

UNIVERSIDADE FEDERAL DO PARANÁ

MARCELO SCARDUELLI

**IDENTIFICAÇÃO E CLONAGEM DE GENES DE CELULASES E
HEMICELULASES DE *Azospirillum brasilense***

**CURITIBA
2014**

MARCELO SCARDUELLI

**IDENTIFICAÇÃO E CLONAGEM DE GENES DE CELULASES E
HEMICELULASES DE *Azospirillum brasilense***

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre pelo programa de pós-graduação em Ciências - Bioquímica, departamento de Bioquímica e Biologia Molecular, do setor de Ciências Biológicas da Universidade Federal do Paraná.

Orientador: Prof. Dr. Emanuel Maltempi de Souza

Co-orientadores: Prof. Dr. Luciano Fernandes Huergo e Dr. Helisson Faoro

CURITIBA

2014

Universidade Federal do Paraná
Sistema de Bibliotecas

Scarduelli, Marcelo

Identificação e clonagem de genes de celulases e hemicelulases de
Azospirillum brasilense. / Marcelo Scarduelli. – Curitiba, 2014.
79 f.: il. ; 30cm.

Orientador: Emanuel Maltempi de Souza

Co-orientadores: Luciano Fernandes Huergo e Helisson Faoro

Dissertação (mestrado) - Universidade Federal do Paraná, Setor de
Ciências Biológicas. Programa de Pós-Graduação em Bioquímica.

1. *Azospirillum brasilense* 2. Celulase I. Título II. Souza, Emanuel
Maltempi de III. Huergo, Luciano Fernandes IV. Faoro, Helisson V.
Universidade Federal do Paraná. Setor de Ciências Biológicas. Programa
de Pós-Graduação em Bioquímica.

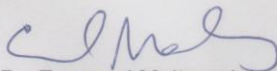
CDD (20. ed.) 589.95

TERMO DE APROVAÇÃO

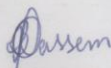
MARCELO SCARDUELLI

**IDENTIFICAÇÃO E CLONAGEM DE GENES DE CELULASES E HEMICELULASES
DE *Azospirillum brasilense***

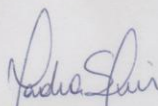
Tese aprovada como requisito parcial para a obtenção do título de Mestre em Ciências-Bioquímica no Programa de Pós-Graduação em Ciências-Bioquímica da Universidade Federal do Paraná, pela Comissão Examinadora formada pelos professores:



Prof. Dr. Emanuel Maltempi de Souza
Departamento de Bioquímica e Biologia Molecular - UFPR



Prof. Dra. Roseli Wassem
Departamento de Genética - UFPR



Dra. Michelle Zibetti Tadra Sfeir
Departamento de Bioquímica e Biologia Molecular - UFPR

AGRADECIMENTOS

Primeiramente agradeço a Deus, pelas oportunidades, proteção e força.

Agradeço aos Professores Emanuel M. de Souza, e Luciano F. Huergo e ao Helisson Faoro pelo apoio, ensinamentos e orientação durante esses 2 anos, estando sempre dispostos a me ajudar, demonstrando muita paciência e comprometimento.

A todos os professores e demais integrantes do Departamento de Bioquímica e Biologia Molecular pelas discussões, dicas e convívio.

À minha esposa, Cleidi Tokarski Scardueli, pelo carinho, amor, apoio, compreensão e ensinamentos. Você torna minha vida muito feliz, te amo muito!

À minha família, em especial aos meus pais Valdir e Graça Scardueli por todo carinho, educação e fé depositada em mim, e aos meus irmãos Bruno e Daniel Scardueli, pelas horas de descontração e apoio. Nós sabemos que mesmo morando em estados diferentes, sempre estamos pertos uns dos outros. Amo muito vocês.

Aos amigos do lab 271 e 272 pelo convívio motivador e animado, em especial para Thiago Estefano Rodrigues, Edileusa Cristina Marques Gerhardt, Maria Isabel Stets, Heloísa Soligo Sanchuki, Luíza Maria de Araújo, Rocio DelPilar Cuaspa Ropaín, Rafaela Perez, Leilane Caline e Vanessa Kessler Chicora.

Aos demais amigos da turma de mestrado 2012.

Aos técnicos Valter de Baura, Roseli Prado e Alexsandro Albani pelo auxílio oferecido constantemente.

À coordenação do curso de pós-graduação em Ciências – Bioquímica, e ao suporte financeiro da CAPES, CNPq e INCT.

“A paz vem de dentro de você mesmo. Não procure à sua volta.”
Buda, Siddhārtha Gautama

RESUMO

Celulases apresentam diversas utilidades industriais, desde uso em etapas de polimento de tecidos até produção de biocombustíveis, especialmente de bioetanol de 2ª geração. Na busca por novas celulases, a prospecção em bactérias colonizadoras de plantas pode levar a identificação de enzimas particulares. Este trabalho teve como objetivo identificar, expressar e caracterizar genes de possíveis celulases e hemicelulases presentes no genoma de *Azospirillum brasilense*, uma bactéria colonizadora de rizosfera de gramíneas. Genes codificando glicosil-hidrolases foram identificados por comparação de sequências. Os genes identificados foram posteriormente clonados e expressos em *E. coli*. Dos 10 genes codificando prováveis glicosil hidrolases, 4 foram clonados em vetor de expressão. Dois dos plasmídeos obtidos, contendo genes de uma celulase e uma hemicelulase, permitiram superexpressar as proteínas recombinantes em *E.coli* que foram purificadas com sucesso. Porém, as 4 enzimas não demonstraram atividade de degradação celulolítica ou hemicelulolítica. Com a finalidade de se identificar se realmente há enzimas com atividade de degradação de celulose ou xilana em *A. brasilense*, a bactéria foi cultivada em meios específicos contendo apenas celulose ou xilana como fonte de carbono. Evidência de crescimento foi observada apenas na presença de xilana, o que leva a crer que há maquinaria ativa capaz de degradar este polímero em *A. brasilense* estirpe FP2.

Palavras-chave: celulases, xilanases, *Azospirillum brasilense*.

ABSTRACT

Cellulases have many industrial uses, from polishing fabrics to biofuel production, especially second generation bioethanol. Searching for new cellulases, bioprospection in plant-root-colonizing bacteria could lead to discover novel enzymes. This study aimed to identify and characterize possible cellulolytic and hemicellulolytic enzymes in genome of *Azospirillum brasilense*. Sequence comparison was used for identification of glycosyl hydrolases (cellulases and xylanases) in the *A. brasilense* draft genome. Four out of 10 selected genes were cloned and overexpressed in *E. coli*. Of these four, a probable cellulase and a hemicellulase were purified successfully. However, the four enzymes had no degradation activity. In order to determine if *A. brasilense* is capable of degrading cellulose or xylan, the bacterium was grown in media containing either cellulose or xylan as sole carbon source. Only xylan sustained *A. brasilense* growth, suggesting the presence of active machinery capable for degrading this polymer in *A. brasilense* strain FP2.

Key-words: cellulases, xylanases, *Azospirillum brasilense*.

LISTA DE FIGURAS

Figura 1: Atuação das três classes principais de celulasas na degradação da celulose em glucose	16
Figura 2: Ilustração estrutural interna de uma biomassa lignocelulósica.....	17
Figura 3: Produtividade média de bioetanol por hectare, para diferentes culturas.....	19
Figura 4: representação esquemática da estrutura de biomassa lignocelulósica após pré-tratamento	20
Figura 5: O ensaio de dinitrossalicilato fundamenta-se na redução do DNS, em meio alcalino, para ác. 3-amino-5-nitrossalicílico.....	39
Figura 6: exemplo de curva padrão para o teste de DNS	40
Figura 7: reação de PCR com 1 M de betaína.....	44
Figura 8: confirmação dos clones transformantes após ligação pTZ57R/T + gene celulaseB	46
Figura 9: Gel de agarose 1% mostrando resultado da restrição com as enzimas <i>EcoRI</i> e <i>HindIII</i> de 3 plasmídeos derivados do vetor pTZ57R/T + inserto celulaseB.....	46
Figura 10: Comparação de sequências do gene CelulaseB (CelB) presente no genoma de <i>A. brasilense</i> Sp245 com o inserto clonado. Programa ClustalW2 (LARKIN <i>et al.</i> , 2007)	49
Figura 11: Confirmação da presença do gene CelulaseB no vetor pET28a – gel de agarose 1%.....	49
Figura 12: Identificação de colônias transformantes contendo o vetor pTZ57R/T com gene XilanaseA	50
Figura 13: Análise de restrição de 3 plasmídeos pTZ57R/T+XilanaseA, utilizando as enzimas <i>EcoRI</i> e <i>HindIII</i> – gel de agarose 1%..	51
Figura 14: Comparação de sequências do gene XilanaseA (XilA) presente no genoma de <i>A. brasilense</i> Sp245 com o inserto clonado	53
Figura 15: confirmação de clones transformantes após ligação pTZ57R/T + gene CelulaseF	54
Figura 16: Análise de restrição com <i>EcoRI</i> e <i>HindIII</i> de 2 plasmídeos derivados da ligação vetor pTZ57R/T + gene CelulaseF, em gel de agarose 1%.....	55

Figura 17: Comparação de sequências do gene CelulaseF (CelF) presente no genoma de <i>A. brasilense</i> Sp245 com o inserto clonado presente em um dos clones obtidos. Programa ClustalW2	57
Figura 18: Comparação de sequências de aminoácidos do gene CelulaseF (CelF) presente no genoma de <i>A. brasilense</i> Sp245 com o inserto clonado	58
Figura 19: Comparação de sequências do gene CelulaseF (CelF) presente no genoma de <i>A. brasilense</i> Sp245 com o segundo inserto clonado	60
Figura 20: Identificação de colônias transformantes contendo o vetor pTZ57R/T + gene CelulaseG.....	61
Figura 21: Análise de restrição com as enzimas <i>EcoRI</i> e <i>HindIII</i> de 3 plasmídeos derivados do vetor pTZ57R/T+CelulaseG, em gel de agarose 1%	62
Figura 22: Perfil eletroforético em gel de poli-acrilamida (10%) de extrato de <i>E. coli</i> BL21 (DE3) expressando as proteínas indicadas (CelB para a CelulaseB, XilA para XilanaseA, CelF para CelulaseF e CelG para CelulaseG)	63
Figura 23: gel de poli-acrilamida (10%) mostrando o resultado do segundo teste de indução feito a 18°C, <i>overnight</i>	64
Figura 24: géis de poli-acrilamida (10%) contendo as amostras obtidas na expressão e purificação da proteína XilanaseA	65
Figura 25: géis de poli-acrilamida (10%) contendo as amostras obtidas na expressão e purificação da proteína CelulaseG.....	66
Figura 26: alinhamento entre sequências similares à proteína CelulaseB utilizando o programa CD-Search.	69
Figura 27: Predição da estrutura secundária da proteína CelulaseB (selvagem e clonada) utilizando o programa PsiPred	70
Figura 28: alinhamento da sequência de XilanaseA obtida e de proteínas similares utilizando o programa CD-Search.	71
Figura 29: Predição da estrutura secundária de XilanaseA (selvagem e clonada) utilizando o programa PsiPred.....	71
Figura 30: alinhamento entre sequências similares à enzima CelulaseF utilizando o programa CD-Search	72
Figura 31: Predição da estrutura secundária para CelulaseF (selvagem e clonada) utilizando o programa PsiPred	73

LISTA DE TABELAS

Tabela 1: Exemplos de celulasas bacterianas.	22
Tabela 2: microrganismos utilizados neste trabalho	25
Tabela 3: plasmídeos usados durante o desenvolvimento deste trabalho.	26
Tabela 4: programa utilizado na técnica de PCR.	31
Tabela 5: iniciadores sintetizados para amplificação e clonagem das glicosil-hidrolases selecionadas de <i>A. brasilense</i> Sp245.	32
Tabela 6: sequência e temperatura de anelamento utilizada para os iniciadores M13 e T7.	33
Tabela 7: soluções utilizadas na eletroforese de proteínas/SDS-PAGE.....	37
Tabela 8: volumes utilizados no preparo dos géis de empilhamento e separação. ...	37
Tabela 9: Genes de prováveis glicosil-hidrolases de <i>A. brasilense</i> Sp245 e características de suas respectivas proteínas.	42
Tabela 10: Identificação da proteína XilanaseA através de espectrometria de massa MS e MS/MS.....	66
Tabela 11: Identificação da proteína CelulaseG através de espectrometria de massa MS/MS.	67
Tabela 12: Identificação de possíveis substratos para as enzimas obtidas	68
Tabela 13: Cultivo de <i>A. brasilense</i> SP245 e FP2 na presença de diferentes polissacarídeos como fonte de carbono.	74

LISTA DE ABREVIATURAS

BGL	beta-glucosidase
CBH	celobiohidrolase
CMC	carboximetilcelulose
	didesoxirribonucleotídeo trifosfato
ddNTP	dimetilsulfóxido
DMSO	3,5-dinitrossalicilato
DNS	desoxirribonucleotídeo trifosfato
dNTP	ditiotreitol
DTT	endoglucanase
EG	ácido α -ciano-4-hidroxicinamico
HCCA	dessorção e ionização a laser assistida
MALDI	por matriz
	espectrometria de massa
MS	espectrometria de massa em
MS/MS	sequência
PA	perssulfato de amônio
Pb	pares de base nitrogenada
PCR	reação em cadeia da polimerase (<i>polymerase chain reaction</i>)
PMF	<i>peptide mass fingerprint</i>
SDS	dodecil-sulfato de sódio
X-GAL	5-bromo-4-cloro-3-indolil- β -D-galactopiranosídeo

ÍNDICE

1	INTRODUÇÃO	14
2	REVISÃO BIBLIOGRÁFICA	15
	2.1 Glicosil-hidrolases	15
	2.2 Processos industriais envolvendo celulasas e hemicelulasas	17
	2.2.1 Hidrólise da biomassa	19
	2.3 As celulasas de <i>trichoderma reesei</i> e de <i>azospirillum sp.</i>	20
3	OBJETIVO	24
	3.1 Objetivos específicos	24
4	MATERIAL E MÉTODOS	25
	4.1 Micro-organismos e vetores	25
	4.2 Meios de cultivo	27
	4.3 Análises de bioinformática	28
	4.4 Manipulação de DNA	29
	4.4.1 Purificação de DNA plasmidial	29
	4.4.2 Purificação de DNA genômico	29
	4.4.3 Amplificação de DNA através do ensaio de PCR	30
	4.4.4 Eletroforese em gel de agarose	33
	4.4.5 Clivagem de DNA com enzima de restrição	33
	4.4.6 Purificação de DNA em gel de agarose	34
	4.4.7 Ligação de DNA	34
	4.4.8 Sequenciamento de DNA	34
	4.5 Transformação bacteriana	35
	4.5.1 Preparo de Células competentes	35
	4.5.2 Transformação	35
	4.6 Manipulação de proteínas	35
	4.6.1 Superexpressão e purificação das proteínas	35

4.6.2 SDS-PAGE - Eletroforese de proteínas sob condição desnaturante.....	36
4.6.3 Dosagem de proteína	38
4.7 determinação atividade de glicosil-hidrolase pelo método do 3,5-dinitrosalicilato	39
5 RESULTADOS E DISCUSSÃO	41
5.1 glicosil-hidrolases de <i>Azospirillum brasilense</i> Sp245.....	41
5.2 Clonagem dos genes de <i>Azospirillum brasilense</i> Sp245.....	43
5.2.1 Clonagem do gene celulaseB de <i>A. brasilense</i> Sp245.....	45
5.2.2 Clonagem do gene XilanaseA de <i>A. brasilense</i> Sp245	50
5.2.3 Clonagem do gene CelulaseF de <i>A. brasilense</i> Sp245	54
5.2.4 Clonagem do gene CelulaseG de <i>A. brasilense</i> Sp245.....	60
5.3 Expressão e purificação dos genes clonados	63
5.4 Ensaios de atividade enzimática.....	67
5.5 Cultivo de <i>A. brasilense</i> em meio contendo apenas cmc e xilana como fonte de carbono	73
6 CONCLUSÃO.....	75
REFERÊNCIAS	76

1 INTRODUÇÃO

As celulasas (ou enzimas celulolíticas) são enzimas capazes de degradar fibras de celulose, o polissacarídeo mais abundante do planeta (ZHANG e LYND, 2006). Juntamente com as hemicelulasas, essas enzimas são vistas como elementos-chave na produção de bioetanol de 2ª geração, apresentando grande interesse industrial.

Empresas já comercializam celulasas para biocombustíveis com atividade satisfatória. Porém, efeitos negativos, como inibição por substrato ou por celobiose e adsorção irreversível à biomassa (CHENG e SUN, 2001), além do custo das enzimas, ainda prejudicam a viabilidade econômica da produção de bioetanol.

Assim, este trabalho tinha como finalidade encontrar novos tipos de celulasas e hemicelulasas que poderiam apresentar atividade satisfatória ou características únicas, permitindo um possível uso em escala industrial.

2 REVISÃO BIBLIOGRÁFICA

2.1 GLICOSIL-HIDROLASES

Enzimas denominadas glicosil-hidrolases são aquelas capazes de hidrolisar ligações glicosídicas em carboidratos. De acordo com a nomenclatura da IUBMB (*International Union of Biochemistry and Molecular Biology*), as glicosil-hidrolases são classificadas como EC 3.2.1.x, onde os três primeiros números representam enzimas capazes de hidrolisar ligações O-glicosídicas e S-glicosídicas. O último número é atribuído de acordo com o tipo de substrato hidrolisado ou o tipo de mecanismo de reação presente na enzima. Esse tipo de classificação pode não ser o adequado para distinguir certas enzimas, principalmente glicosil-hidrolases. HENRISSAT (1991) percebeu que esse tipo de classificação não evidencia propriedades estruturais específicas de certas enzimas, deixando de distinguir enzimas que sofreram evolução divergente (mudança na especificidade enzima-substrato) ou convergente (polipeptídeos com diferenças estruturais catalisando a mesma reação no mesmo substrato). Além disso, o autor notou que enzimas que degradassem vários tipos de substratos mostrar-se-iam difíceis de serem classificadas. Dessa forma, o autor criou a classificação das glicosil-hidrolases por famílias, baseada na similaridade de sequência de aminoácidos, e, conseqüentemente, na similaridade de suas estruturas proteicas.

Atualmente, o banco de dados CAZY (do inglês, *Carbohydrate-Active enZymes database*) apresenta uma ótima fonte de informação sobre as variadas famílias de glicosil-hidrolases, que vão da família 1 (enzimas com estrutura de barril (β/α)₈), até a família 133 (enzimas bifuncionais). Dentro desta classificação, são encontradas celulasas e hemicelulasas em diversas famílias.

As celulasas são enzimas que degradam a celulose, um homopolímero formado por unidades de D-glucose $\beta(1\rightarrow4)$ ligadas. Para ocorrer essa degradação, há a necessidade de se utilizar principalmente 3 tipos de enzimas: endoglucanases (EG), exoglucanases ou celobiohidrolases (CBH) e β -glucosidases (BGL) (COUGHLAN e LJUNGDAHL, 1988). A figura 1 ilustra a sinergia existente entre essas classes de enzimas.

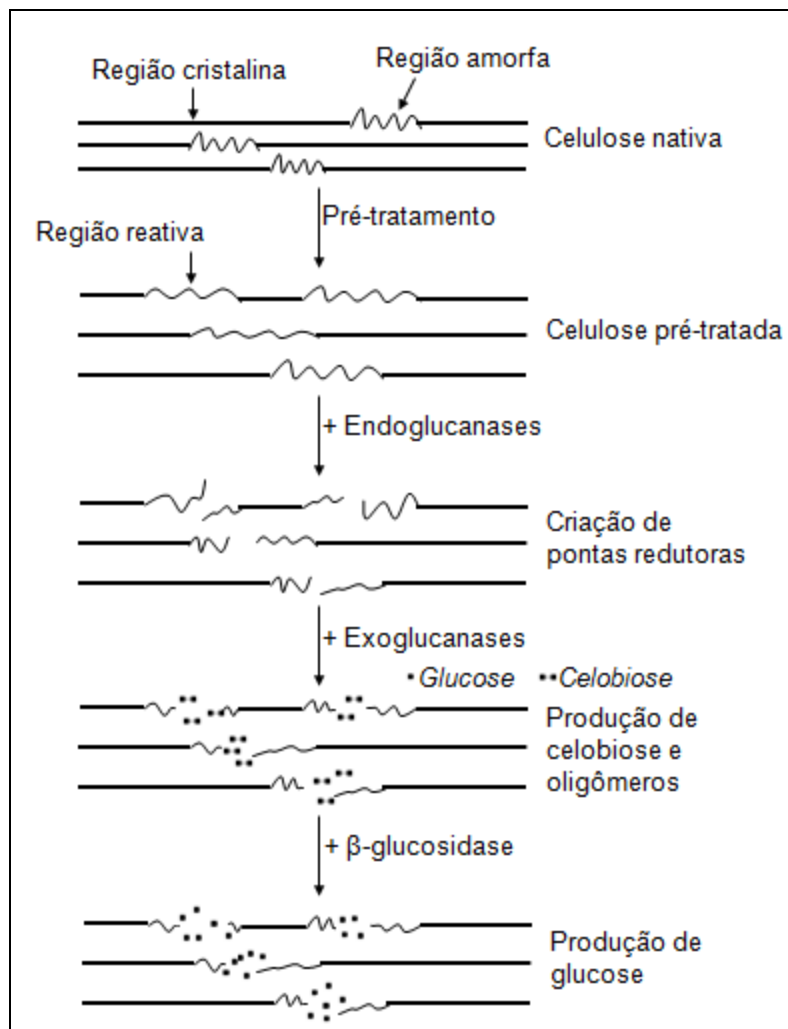


Figura 1: Atuação das três classes principais de celulasas na degradação da celulose em glucose (TAHERZADEH e KARIMI, 2007 – adaptado).

As EGs (EC 3.2.1.4) atuam hidrolisando ligações glicosídicas presentes em regiões amorfas da celulose, expondo novas pontas livres nas fibras. Na classificação de Henrissat, há representantes dessas enzimas nas famílias 5, 6, 9, 12, 44, 45, 51, 74 e 124.

Já as CBHs (EC 3.2.1.91) iniciam a degradação da celulose a partir das pontas livres recém-formadas, liberando unidades de celobiose. Essas enzimas ainda podem ser classificadas em CBH I (atuam sobre terminais redutores da celulose) e CBH II (clivam terminais não redutores da fibra). Elas são encontradas nas famílias 6, 7, 9, 48 e 74.

Por fim, as BGLs (EC 3.2.1.21) finalizam a degradação da fibra celulósica, hidrolisando as celobioses formadas. Essas enzimas estão presentes nas famílias 1, 3, 4, 5, 9, 13, 17, 30, 31, 63, 97, 116, 122 e 133.

Durante o processo de degradação da celulose a partir de um material lignocelulósico (biomassa), algumas enzimas auxiliares podem ser utilizadas, como as hemicelulases. Esse grupo de enzimas engloba catalisadores capazes de degradar a hemicelulose (figura 2), uma estrutura amorfa formada por xiloglucanas, xilanas, mananas e glucomananas, cuja principal função é dar rigidez para a parede celular vegetal através da interação com a celulose e com a lignina (complexo formado por alcoóis aromáticos) (SCHELLER e ULVSKOV, 2010). Assim, as xilanases, glucuronidases, acetil-esterases, β -xilosidases, galactomananases e glucomananases podem auxiliar nesse processo de degradação, expondo as fibras de celulose. Essas enzimas são encontradas em aproximadamente 25 famílias de glicosil-hidrolases.

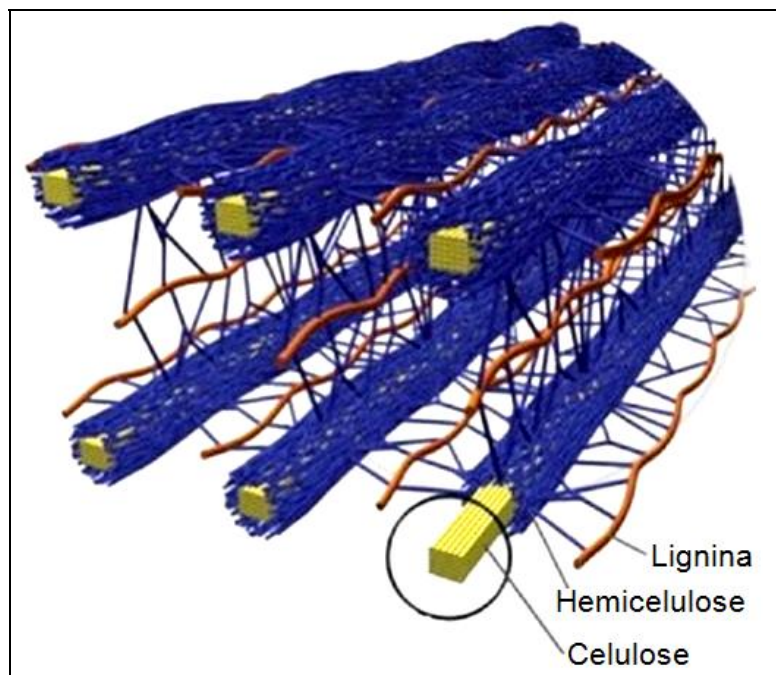


Figura 2: Ilustração estrutural interna de uma biomassa lignocelulósica. Fonte: INTECH – adaptado.

2.2 PROCESSOS INDUSTRIAIS ENVOLVENDO CELULASES E HEMICELULASES

Celulases atraem muito interesse em pesquisas devido à sua elevada possibilidade de usos em processos industriais, entre eles: remoção de cor e biopolimento em têxteis, branqueador em lava-roupas, diminuição de viscosidade em bebidas, além de seu uso na conversão de biomassa em biocombustíveis (PANG *et*

al., 2009). Devido a essa variedade de aplicação, há a necessidade de identificar uma variedade de enzimas que possam se adaptar em cada condição do processo industrial.

Antigamente, o processo de remoção de cor na indústria têxtil era feito apenas utilizando pedras-pomes juntamente com o tecido a ser desgastado, durante etapas de lavagem. Esse efeito é obtido através de choques mecânicos contra as fibras do tecido, gerado pelas pedras, ocorrendo a quebra das fibras celulósicas e posterior liberação do corante. Agora, as indústrias utilizam celulasas nesse processo, que substituem as etapas de lavagem com as pedras e consegue gerar o mesmo efeito final de forma mais facilitada (GALANTE *et al.*, 1998).

Já o processo de biopolimento consiste em eliminar as microfibrilas superficiais da fibra de algodão recém produzida, conferindo um aspecto mais liso, brilhoso e macio ao tecido (MIETTINEN-OINONEN, *et al.*, 2005). Essa técnica utiliza celulasas e também alguns tipos de proteases.

Quanto ao efeito branqueador, algumas celulasas específicas, adicionadas ao sabão em pó, também atuam retirando as microfibrilas que se soltaram das fibras principais do tecido. Esse descolamento ocorre após sucessivas lavagens, fazendo com que a luz refletida nessas microfibrilas seja dispersa, o que causa a aparência gasta ao tecido. Além disso, algumas manchas particuladas podem estar agregadas nessas microfibrilas soltas, criando o aspecto “encardido” à roupa. Como as celulasas conseguem retirar essas microfibrilas, as manchas também podem ser retiradas, resultando em um tecido mais branco (NOVOZYMES, 2003).

Na indústria de bebidas, podem-se utilizar celulasas e hemicelulasas para diminuir a viscosidade de misturas. Durante a produção de cerveja, a cevada pode gerar um mosto com excesso de polissacarídeos não-amílicos, constituídos principalmente por β -glucanas solúveis. Esses polissacarídeos formam géis que dificultam a filtração da mistura e aumentam a turbidez do produto final, prejudicando sua qualidade. Para contornar esses problemas, utilizam-se β -glucanases nas etapas de trituração ou fermentação da cevada. Além disso, utilizam-se enzimas também na produção de vinho. Pectinases, β -glucanases e hemicelulasas facilitam as etapas de maceração da casca da uva, de clarificação e de filtração do mosto, melhorando a qualidade e estabilidade do vinho (GALANTE *et al.*, 1998).

2.2.1 Hidrólise da biomassa

Denomina-se biomassa qualquer recurso natural que dispõe de energia capaz de ser processada fornecendo formas bioenergéticas mais elaboradas e adequadas, como os biocombustíveis: bioetanol e biodiesel, por exemplo (BNDES e CGEE, 2008). Vários tipos de biomassa podem ser utilizados na produção de bioetanol, como a cana-de-açúcar, milho, mandioca, beterraba, sorgo, trigo. Porém, quando se analisa o potencial de produtividade, a cana-de-açúcar se destaca (figura 3).

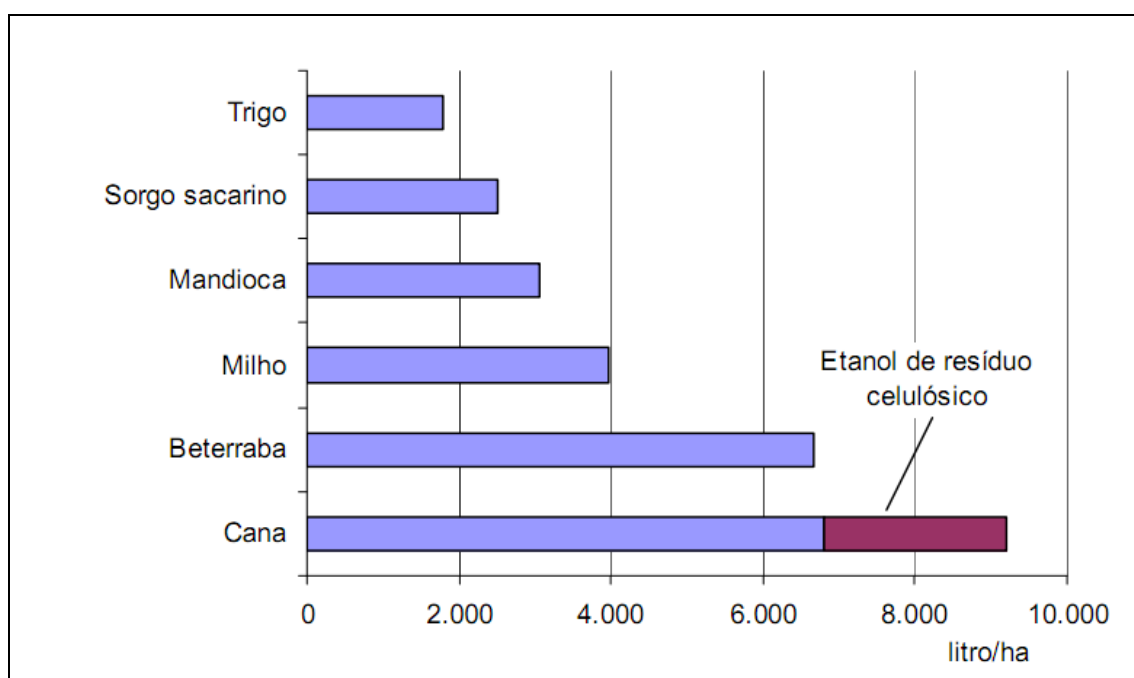


Figura 3: Produtividade média de bioetanol por hectare, para diferentes culturas (BNDES e CGEE, 2008).

Além disso, o bagaço de cana-de-açúcar é uma matéria-prima com baixo valor comercial, facilitando seu uso na produção de bioetanol (BNDES e CGEE, 2008). Esse resíduo pertence ao grupo das biomassas lignocelulósicas, que apresentam como constituintes principais a celulose, a hemicelulose e a lignina (figura 2). A conformação desses componentes ocorre de maneira complexa na biomassa *in natura*, tornando-a recalcitrante, ou seja, o acesso por enzimas à celulose se encontra dificultada. Por isso, é indispensável a utilização de um processo de pré-tratamento para viabilizar a posterior hidrólise enzimática.

O caráter altamente recalcitrante da lignocelulose nativa pode ser sensivelmente atenuado pela explosão a vapor (figura 4). Essa técnica consiste em

aquecer a biomassa a altas temperaturas e pressão, seguido de súbita diminuição da pressão no reator, ocorrendo o fenômeno de explosão das fibras (BABCOCK, 1932). O efeito causado sobre a organização estrutural da celulose aumenta consideravelmente sua área superficial e, por conseguinte, sua susceptibilidade à hidrólise enzimática. Isso promove a redução da quantidade de enzima a ser adicionada e, consequentemente, redução de custos (FOODY e FOODY, 1991).

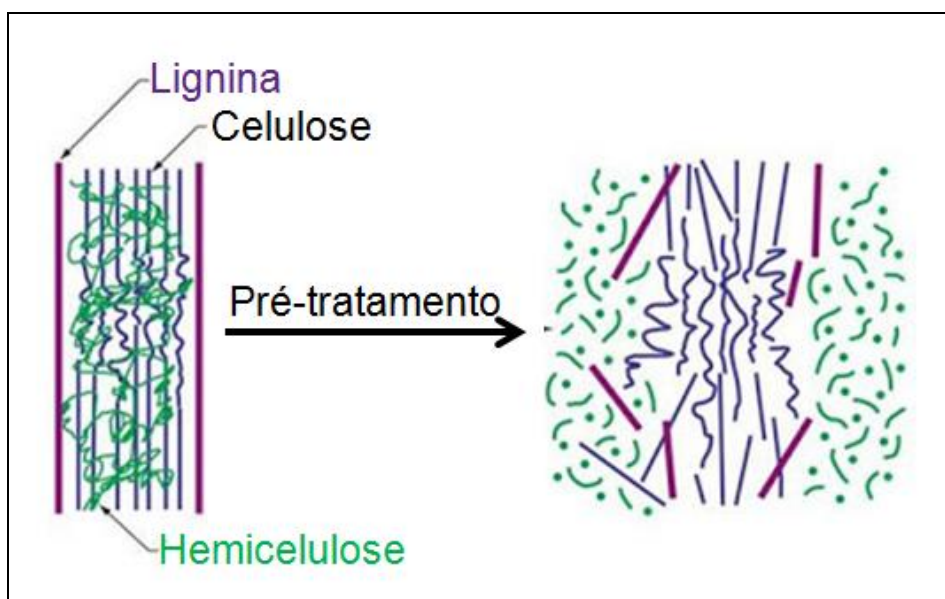


Figura 4: representação esquemática da estrutura de biomassa lignocelulósica após pré-tratamento. Fonte: <http://biofueloutlook.blogspot.com.br/>

A hidrólise da biomassa pode ser feita tanto química (utilizando ácido sulfúrico como catalisador) quanto enzimaticamente, porém estudos mostram que o uso da hidrólise enzimática é mais vantajoso que o método químico em vários aspectos: maior taxa de eficiência celulolítica, ausência de perda de substrato por modificações químicas e condições físico-químicas de processo brandas e não-corrosivas (temperaturas menores de reação, pH próximo ao neutro), além de utilizar reagentes biodegradáveis e atóxicos (BOM e FERRARA, 2007).

2.3 AS CELULASES DE *TRICHODERMA REESEI* E DE *AZOSPIRILLUM SP.*

O fungo *Trichoderma reesei* é o organismo celulolítico (degradador de celulose) mais estudado atualmente. Seu complexo celulolítico foi utilizado como base na formulação do modelo de biodegradação da celulose. O gênero

Trichoderma compreende fungos filamentosos da classe Ascomycetes (ascomicetos). Por apresentarem a habilidade de produzir hidrolases extracelulares em grande quantidade (pode chegar a 30g de celulases por litro de cultivo), essas espécies apresentam muito interesse industrial. Suas enzimas são utilizadas na indústria têxtil, em alimentos e na indústria de papel (MIETTINEN-OINONEN, 2004).

Dentre as enzimas secretadas, a espécie *T. reesei* produz EGs e CBHs associadas a estruturas não catalíticas denominadas CBMs (módulos de ligação a carboidrato, do inglês *carbohydrate-binding modules*), que permitem a ligação das enzimas ao substrato. A ausência desses módulos não prejudica a atividade sobre substratos solúveis, mas diminui consideravelmente a hidrólise de substratos insolúveis, como as porções cristalinas da celulose (ARAI *et al.*, 2003).

Além de fungos, diversas bactérias também apresentam celulases em seu genoma (tabela 1). Já foram identificadas enzimas com atividade ótima em temperaturas variando de 40 a 74°C e pH indo de 4,8 a 11. Essa flexibilidade de condições de hidrólise viabiliza (e estimula) a utilização de celulases bacterianas em diversos processos industriais.

Tabela 1: Exemplos de celulases bacterianas (SADHU e MAITI, 2013 – adaptado).

Espécie	Enzima	Peso Molecular	Temperatura ótima (°C)	pH ótimo
Aeróbicas				
<i>Bacillus licheniformis</i>	Endoglucanase	-	55	6,1
<i>Bacillus sp</i> (alcalofílico)	Endoglucanase	54	-	5,0-11,0
<i>Bacillus sp</i> (neutrofílico)	Endoglucanase	35	50	7,0-10,0
<i>Bacillus subtilis</i> DLG	Endoglucanase	35	55	4,8
<i>Cellulomonas uda</i>	Exocelobiohidrolase	81	45-50	5,5-6,5
<i>Microbispora bispora</i>	Endoglucanase	44	-	5,5-7,2
	Exoglucanase	75	-	5,9-7,2
<i>Thermomonospora fusca</i> YX	Endoglucanase1	94	74	6,0
	Endoglucanase2	46	58	6,0
Anaeróbicas				
<i>Bacteroides succinogenes</i>	Endoglucanase	118	39	5,8
<i>Clostridium josui</i>	Endoglucanase	45	60	6,8
<i>Clostridium thermocellum</i>	Endoglucanase	94	62	5,2
<i>Clostridium thermocopriae</i>	Endoglucanase	46	-	6,5

Dentre as fontes bacterianas, podemos destacar a espécie *Azospirillum brasilense*. Bactérias pertencentes ao gênero *Azospirillum* são encontradas principalmente em ambientes terrestres, onde podem estar associadas a plantas,

promovendo crescimento vegetal através da fixação de nitrogênio e secreção de fito-hormônios (STEENHOUDT e VANDERLEYDEN, 2000). Essas bactérias são pertencentes à classe Alpharotobacteria, são Gram negativas, curvas, móveis e presentes em várias regiões do planeta (YOUNG, 1992), sendo a espécie *A. brasilense* a mais estudada (HUERGO, 2006).

Como essa bactéria é capaz de colonizar raízes de plantas, seu genoma deve apresentar enzimas que permitam sua entrada através da estrutura vegetal. Análise do genoma da estirpe Sp245 de *A. brasilense* levou a identificação de 15 genes codificando para prováveis glicosil-hidrolases (WISNIEWSKI-DYE *et al.*, 2011). Ainda é desconhecido como ocorre a invasão de *A. brasilense* Sp245 na planta, porém acredita-se que essa colonização ocorra realmente por degradação enzimática, tanto pectinolítica quanto celulolítica (MEHDIPOUR-MOGHADDAM *et al.*, 2010).

3 OBJETIVO

O objetivo geral deste projeto foi identificar e caracterizar novas enzimas no genoma de *Azospirillum brasilense*.

3.1 OBJETIVOS ESPECÍFICOS

- Identificar *in silico* genes que codificam celulasas e xilanases em sequência genômica de *A. brasilense*;
- Amplificar genes de interesse a partir de iniciadores específicos;
- Sequenciar insertos de DNA de clones e/ou de produtos de PCR referentes a genes que codifiquem glicosil-hidrolases;
- Clonar e expressar genes de glicosil-hidrolases selecionados em *E. coli*;
- Purificar e caracterizar as celulasas/hemicelulasas expressas.

4 MATERIAL E MÉTODOS

4.1 MICRO-ORGANISMOS E VETORES

Os microrganismos e vetores utilizados neste trabalho estão descritos na tabela 2 e 3. O genoma de *A. brasilense* Sp245 foi obtido no *GenBank* (bioproject: PRJEA162161, ID: 162161).

Tabela 2: microrganismos utilizados neste trabalho.

Microrganismo	Características	Referência
<i>Escherichia coli</i> BL21 (λDE3)	hsdS _B , gal(λclts 857 ind1 Sam7 nin5 lacUV5-T7 gene 1)	SAMBROOK <i>et al.</i> , 1989
<i>Escherichia coli</i> DH10B	Sm ^r , F ⁺ [proAB ⁺ lacZΔM15]	SAMBROOK <i>et al.</i> , 1989
<i>Azospirillum</i> <i>brasilense</i> Sp245	Estirpe selvagem	BALDANI, BALDANI e DÖBEREINER, 1983
<i>Azospirillum</i> <i>brasilense</i> FP2	NaI ^r Sm ^r Estirpe selvagem SP7 Nif ⁺	PEDROSA e YATES, 1984

Tabela 3: plasmídeos usados durante o desenvolvimento deste trabalho.

Plasmídeo	Características	Referência
pTZ57R/T	Amp ^r , vetor de clonagem	Fermentas
pET28a	Km ^r , vetor de expressão (promotor T7)	Novagen
pMS_zCelB	Amp ^r , gene AZOBR_p140078 clonado em pTZ57R/T (com mutações L138P e N437D)	Este trabalho
pMS_CelB	Km ^r , gene AZOBR_p140078 clonado em pET28a (com mutações L138P e N437D)	Este trabalho
pMS_zXilA	Amp ^r , gene AZOBR_p1110166 clonado em pTZ57R/T (com mutações A41V e L56P)	Este trabalho
pMS_XilA	Km ^r , gene AZOBR_p1110166 clonado em pET28a (com mutações A41V e L56P)	Este trabalho
pMS_zCelF	Amp ^r , gene AZOBR_150049 clonado em pTZ57R/T (com mutações K147E e D268G)	Este trabalho
pMS_CelF	Km ^r , gene AZOBR_150049 clonado em pET28a (com mutações K147E e D268G)	Este trabalho
pMS_zCelG	Amp ^r , gene AZOBR_p470008 clonado em pTZ57R/T	Este trabalho
pMS_CelG	Km ^r , gene AZOBR_p470008 clonado em pET28a	Este trabalho

4.2 MEIOS DE CULTIVO

Os meios de cultura utilizados para as estirpes de *E. coli* foram os meios LA e LB (SAMBROOK *et al.*, 1989) acrescidos de:

- Ampicilina (250 µg/mL) e X-Gal (5-bromo-4-cloro-3-indolil-β-D-galactopiranosídeo, 250 µg/mL), durante etapas de clonagem com o vetor pTZ57R/T;
- Canamicina (50 µg/mL), durante clonagem com vetor pET28a e superexpressão de proteínas.

- Meio LB (SAMBROOK *et al.*, 1989):

Extrato de levedura: 5,0 g/L

NaCl: 10,0 g/L

Triptona: 10,0 g/L

- Meio LA (SAMBROOK *et al.*, 1989):

Extrato de levedura: 5,0 g/L

NaCl: 10,0 g/L

Triptona: 10,0 g/L

Ágar bacteriológico: 15,0 g/L

As estirpes de *E. coli* foram cultivadas a 37°C. As culturas em meio líquido foram incubadas em agitador rotatório a 120rpm. A preservação das estirpes e dos clones gerados foi feita em glicerol 50% e estocadas a -20°C.

Para o cultivo de *A. brasilense*, utilizou-se o meio NFb-lactato (MACHADO *et al.*, 1995), contendo antibióticos estreptomicina (20 µg/mL) e ácido nalidíxico (5 µg/mL). Como fonte de nitrogênio, foi utilizado 20 mM de NH₄Cl. Para meio sólido, adicionou-se 15 g/L de ágar. Após semeadura, o meio líquido foi mantido a 30°C a 120 rpm, em agitador, e o meio sólido, em estufa a 30°C.

- Meio Nfb-lactato (MACHADO *et al.*, 1995):

KH₂PO₄: 4,0 g/L

K₂HPO₄: 6,0 g/L

MgSO₄.7H₂O: 0,2 g/L

NaCl: 0,1 g/L

CaCl₂: 0,02 g/L

Ácido nitrilo triacético: 0,056 g/L

FeSO₄.7H₂O: 0,02 g/L

Lactato de sódio: 5,0 g/L

Biotina: 1,0.10⁻⁴ g/L

Na₂MoO₄.2H₂O: 2,0.10⁻³ g/L

MnSO₄.H₂O: 2,35 .10⁻³ g/L

H₃BO₃: 2,8.10⁻³ g/L

CuSO₄.5H₂O: 8,0.10⁻⁵ g/L

ZnSO₄.7H₂O: 2,4.10⁻⁴ g/L

O meio utilizado para avaliar se as estirpes de *A. brasilense* (Sp245 e FP2) conseguiam crescer em cultura contendo apenas CMC (carboximetilcelulose) ou xilana como única fonte de carbono foi denominado meio TESTE. Este meio foi modificado por Ana Paula Andreazza a partir dos meios descritos por YAN e colaboradores (2013), AMORE e colaboradores (2012) e HANKIN e ANAGNOSTAKIS (1976).

- Meio TESTE (Ana Paula Andreazza):

K ₂ HPO ₄ : 1,6 g/L	Extrato de levedura: 0,5 g/L
KH ₂ PO ₄ : 0,2 g/L	CMC: 5 g/L OU xilana: 3 g/L
(NH ₄) ₂ SO ₄ : 1 g/L	Na ₂ MoO ₄ .2H ₂ O: 0,01 g/L
MgSO ₄ .7H ₂ O: 0,2 g/L	MnSO ₄ .H ₂ O: 1,2 .10 ⁻² g/L
FeSO ₄ .7H ₂ O: 0,01 g/L	H ₃ BO ₃ : 1,4.10 ⁻² g/L
Lactato de sódio: 5,0 g/L	CuSO ₄ .5H ₂ O: 4,0.10 ⁻⁴ g/L
NaCl: 0,1 g/L	ZnSO ₄ .7H ₂ O: 1,2.10 ⁻³ g/L
CaCl ₂ .2H ₂ O: 0,02 g/L	

A xilana utilizada é originada de madeira de faia (*beechwood*), da Sigma, composta por cadeia linear de poli(β-D-xilopirranose[1→4]). O CMC (carboximetilcelulose sódica) foi produzido por Minérios do Brasil e é caracterizado como CMC de alta viscosidade (não apresenta informações de origem e estrutura).

As condições de cultivo para este meio foi as mesmas utilizadas para o meio líquido NfbHPN-lactato (3 dias a 30°C, em estufa), contendo antibióticos estreptomicina (20 µg/mL) e ácido nalidíxico (5 µg/mL).

4.3 ANÁLISES DE BIOINFORMÁTICA

O programa BlastX (ALTSCHUL *et al.*, 1997) foi utilizado para se obter índices de similaridade e identidade entre os genes estudados com proteínas presentes no banco de dados GenBank (<http://www.ncbi.nlm.nih.gov>). Já o programa ClustalW2 (LARKIN *et al.*, 2007) foi usado para gerar alinhamentos entre sequências, permitindo a avaliação da conservação de bases ou aminoácidos. O ProtParam

(GASTEIGER *et al.*, 2005) permite a obtenção de parâmetros físico-químicos (teóricos) a partir de uma sequência de aminoácidos. Para identificar domínios conservados, famílias de proteínas e sítios funcionais foi utilizado o programa ScanProsite (SIGRIST *et al.*, 2002). O programa PSORT (NAKAI e KANEHISA, 1991) foi utilizado na predição de localização subcelular de proteínas. O CD-SEARCH (MARCHLER-BAUER *et al.*, 2011) permite identificar domínios conservados através de comparação entre uma sequência de interesse com o banco de dados GenBank. O programa permite também comparar a sequência submetida com outras sequências similares, permitindo a identificação de resíduos conservados. O programa PsiPred (JONES, 1999) foi utilizado na predição de estruturas secundárias.

4.4 MANIPULAÇÃO DE DNA

4.4.1 Purificação de DNA plasmidial

A obtenção de DNA plasmidial foi realizada pelo método de lise alcalina (SAMBROOK *et al.*, 1989).

4.4.2 Purificação de DNA genômico

A purificação do DNA genômico de *Azospirillum brasilense* foi feito de maneira similar ao método aplicado a *Herbaspirillum spp.* (SOUZA, 1990).

A partir de 10 mL de cultivo saturado de *A. brasilense* (em meio NfbHPN-lactato, mais antibióticos), as bactérias foram coletadas por centrifugação a 5000 rpm por 10 minutos. Depois de descartar o sobrenadante, as células foram lavadas com 10 mL de tampão GET (50 mM glucose, 25 mM Tris-HCl pH 8,0 e 10 mM de EDTA pH 8,0) 2 ou 3 vezes. O pellet resultante foi ressuspenso em 500 µL de tampão GET, adicionado 2 µL de lisozima (10 µg/mL) e deixado a 30°C durante 5 horas. Em seguida foram adicionados 50 µL de SDS (dodecil sulfato de sódio) 10% e incubado por 5 minutos a temperatura ambiente. Ao lisado foram acrescentados 200 µg/mL de pronase e a mistura foi incubada a 37°C por 16 horas. A preparação

foi desproteínada pela adição de 500 µL de fenol equilibrado com tampão Tris-HCl (pH 8,0), depois foi feita a homogeneização e centrifugada a 13400 rpm a 15 minutos. Uma segunda etapa de desproteínização foi realizada com 500 µL de fenol-clorofórmio, seguida de uma etapa com 500 µL de clorofórmio-álcool isoamílico (24:1). Essas etapas de desproteínização podem ser repetidas até não ocorrer mais a formação de “material esbranquiçado” (precipitação protéica) na interface aquosa-orgânica. O DNA resultante na fase aquosa foi então precipitado pela adição de 0,6 volume de isopropanol. Após centrifugação (13400 rpm, 15 minutos), o isopropanol foi descartado e o precipitado foi lavado com 1 mL de etanol 70%, e precipitado foi seco em bomba de vácuo ou em estufa a 37°C. Quando seco, o DNA foi dissolvido em 100 µL de água ultrapura autoclavada.

4.4.3 Amplificação de DNA através do ensaio de PCR

Para a obtenção dos genes estudados foi utilizado o sistema de reação em cadeia da polimerase (PCR) (KOCHER e WILSON, 1991). Os iniciadores utilizados neste trabalho foram sintetizados pela IDT. As condições padrão do ensaio de PCR foram: 10 mM Tris-HCl pH 8,8, 50 mM KCl, 1 mM MgCl₂, 0,3 mM de dNTP, 10 pmol de cada iniciador (*forward* e *reverse*), aproximadamente 10 ng de DNA molde e 1 U de Taq DNA polimerase (purificada pela Professora Leda Chubatsu) em 25 µL de reação. Foi utilizado 1 M de betaína como auxiliar de amplificação (maiores detalhes na seção de resultados). O programa utilizado no termociclador (Veriti, Applied Biosystems) e os iniciadores utilizados estão mostrados nas tabelas 4 e 5, respectivamente.

Tabela 4: programa utilizado nas amplificações de DNA através da técnica de PCR.

N° de repetições	Temperatura	Tempo de incubação
1	96°C	2 min
25 (ciclo)	94°C	30 seg
	Tm*	50 seg
	72°C	2 min e 30 seg
1	72°C	5 min

Tm*: temperatura de anelamento variável, de acordo com o iniciador utilizado.

Tabela 5: iniciadores sintetizados para amplificação e clonagem das glicosil-hidrolases selecionadas de *A. brasilense* Sp245.

Nº	Gene	Iniciadores	Tm (°C)*	Amplificado esperado (pb)	Sítios de clivagem
1	AZOBR_p1110166	Fwd: 5' ACG CAT ATG GCC CAG ACC TC Rev: 5' CTG TCG GGG ATC CTT ACA GG	51	1379	<i>NdeI</i> <i>BamHI</i>
2	AZOBR_p310014	Fwd: 5' CTG TCA TAT GAG AGC GCA G Rev: 5' AGC CTC GAG AAC CAT GCC TG	49	1168	<i>NdeI</i> <i>XhoI</i>
3	AZOBR_p440082	Fwd: 5' ACC ACA TAT GAC CAC GAC GC Rev: 5' CCT CGA GGC TCA CCC CTA CA	47	840	<i>NdeI</i> <i>XhoI</i>
4	AZOBR_p140078	Fwd: 5' AAG AAT TCC ATA TGC GGC GG Rev: 5' CTG GAT CCG AGC CTC TCA AG	45	1475	<i>NdeI</i> <i>BamHI</i>
5	AZOBR_p280124	Fwd: 5' AGT CCC CCA TAT GGC ACT CT Rev: 5' ATG GAT CCG AAC AGC CTG CT	47	1375	<i>NdeI</i> <i>BamHI</i>
6	AZOBR_p280134	Fwd: 5' TGA GGC ATA TGG AGA CCA CG Rev: 5' GGC AGG ATC CGG AAT TGT CC	49	1435	<i>NdeI</i> <i>BamHI</i>
7	AZOBR_p280136	Fwd: 5' GAT ATG GGG CAT ATG CGC GG Rev: 5' GCT TCG AGG ATC CAA TCG CC	47	1169	<i>NdeI</i> <i>BamHI</i>
8	AZOBR_150049	Fwd: 5' GGC TGA GAC ATA TGG GCA TC Rev: 5' GGA TCC TAA GGC TCC CGA TC	47	1354	<i>NdeI</i> <i>BamHI</i>
9	AZOBR_p470008	Fwd: 5' AGA CAT ATG GCG ACG ACC TC Rev: 5' AGG ATC CCG GGC GAT TTC TC	49	2112	<i>NdeI</i> <i>BamHI</i>
10	AZOBR_p1110164	Fwd: 5' CTT CAT ATG GCA ACG ACC AC Rev: 5' GGA CCT TCA CCG CTC TCG AG	50	2500	<i>NdeI</i> <i>XhoI</i>

* temperatura de anelamento para o par de iniciadores utilizados nos ensaios de PCR.

Os DNA moldes utilizados foram DNA genômico de *Azospirillum brasilense* estirpe Sp245 e colônias transformantes (*E. coli* DH10 e BL21 DE3). Para os ensaios de PCR utilizando colônias, foram utilizados os iniciadores M13 universal e reverso (para o vetor pTZ57R/T) e promotor e terminador T7 (para o vetor pET28a), cujas sequências estão mostradas na tabela 6.

Tabela 6: sequência e temperatura de anelamento utilizada para os iniciadores M13 e T7.

Gene	Iniciadores	T _m (°C)*
M13	Universal: 5' TGT AAA ACG ACG GCC AGT	50
	Reverse: 5' CAG GAA ACA GCT ATG ACC	
	Promotor: 5' TAA TAC GAC TCA CTA TAG GG	
T7	Terminador: 5' TAA TAC GAC TCA CTA TAG GG	50

* temperatura de anelamento utilizados nos ensaios de PCR

4.4.4 Eletroforese em gel de agarose

A eletroforese de DNA foi feita em gel de agarose (1 a 2% m/v) horizontal conforme descrito por SAMBROOK e colaboradores (1989). Para visualização do DNA, utilizou-se brometo de etídeo a 0,5 µg/mL e trans-iluminador de luz ultravioleta. O perfil eletroforético foi registrado em equipamento UVP BioImaging Systems (sistema de vídeo-imagem acoplado).

4.4.5 Clivagem de DNA com enzima de restrição

As endonucleases de restrição utilizadas (*Bam*HI, *Nde*I, *Hind*III, *Eco*RI e *Xho*I) foram adquiridas da New England Biolabs ou Fermentas. As condições de reação foram aquelas sugeridas pelo fabricante.

4.4.6 Purificação de DNA em gel de agarose

O kit de purificação “QIAquick Gel Extraction Kit”, da QIAGEN, foi utilizado para extrair DNA diretamente de gel de agarose após corrida eletroforética. O protocolo usado foi o recomendado pelo fabricante.

4.4.7 Ligação de DNA

O DNA plasmidial linearizado (vetor) foi misturado ao DNA inserto, na razão molar 1:5, respectivamente, em tampão de ligação (50 mM Tris-HCl pH 7,5; 10 mM MgCl₂, 10 mM ATP e 10 mM DTT). Adicionou-se ao sistema 0,5 a 2U de T4 DNA ligase, que foi incubado a 16°C por pelo menos 16 horas.

4.4.8 Sequenciamento de DNA

O sequenciamento de DNA foi realizado pelo método de SANGER *et al.* (1977), conhecido como terminação de cadeia utilizando didesoxirribonucleotídeos (ddNTPs), modificados para sequenciador automático de DNA (Applied Biosystems modelo 3500XL) e didesoxirribonucleotídeos fluorescentes (*dye terminators*). As reações de sequenciamento foram realizadas utilizando o sistema *BigDye* (Applied Biosystems). Os parâmetros de amplificação foram definidos de acordo com os iniciadores utilizados.

Após a amplificação, foi adicionado ao sistema 10 µL de água ultrapura, 2,0 µL de acetato de amônio 5,0 M e 66 µL de etanol absoluto. O DNA precipitado foi recuperado por centrifugação a 13000 rpm por 45 minutos. O sobrenadante foi descartado e o precipitado foi lavado com 100 µL de etanol 70%. Após o descarte do sobrenadante, o produto da reação foi secado em fluxo laminar e analisado em sequenciador automático ABI XL3500.

4.5 TRANSFORMAÇÃO BACTERIANA

4.5.1 Preparo de Células competentes

Algumas colônias de *E. coli* crescidas em meio LA foram inoculadas em 5 mL de meio LB na presença de antibióticos se necessário. A cultura foi incubada por 14-16 horas, sob agitação a 37°C. Todo o pré-inóculo foi transferido para um erlenmeyer de 1 L contendo 250 mL de LB mais antibióticos e a cultura foi mantida a 37°C até atingir absorbância a 600 nm entre 0,3 e 0,4. As células permaneceram 30 minutos no gelo, então centrifugadas a 3.000 g por 10 minutos a 4°C. O *pellet* foi ressuspenso em 30 mL de uma solução gelada de CaCl₂ 0,1 M. Após 1 hora no gelo, as células foram novamente centrifugadas e ressuspensas em uma solução gelada de CaCl₂ 0,1 M e 15% de glicerol. Frações de 200 µL foram aliquotadas, armazenadas em tubos de 1,5 mL a -80°C.

4.5.2 Transformação

As alíquotas de células foram retiradas do congelador -80° C e mantidas em gelo por 30 minutos. Quantidades suficientes de DNA (de 20 a 50 ng de plasmídeo ou sistema de ligação) foram adicionados às células competentes, incubadas em banho de gelo por mais 30 minutos, seguido de 2 minutos em banho maria a 42°C e voltando para o gelo por mais 10 minutos. Depois, adicionou-se 1 mL de LB às células, que foram mantidas a 37°C por 40 minutos. Cem microlitros da suspensão de células foram plaqueados em LA com antibióticos adequados.

4.6 MANIPULAÇÃO DE PROTEÍNAS

4.6.1 Superexpressão e purificação das proteínas

A purificação foi feita a partir de células de *E. coli* da estirpe BL21 (λDE3) contendo o plasmídeo específico, que expressa a proteína de interesse. Foram inoculadas colônias transformantes em 20 mL de LB mais antibióticos que foi

incubado durante a noite a 37°C. Esse pré-inóculo foi vertido em um erlenmeyer de 2L contendo 800 mL de LB com antibióticos, e incubado em um agitador rotatório a 120 rpm e 37°C até atingir absorbância a 600 nm entre 0,6 a 0,7. Logo depois, a cultura foi mantida a 37°C ou transferida para outro agitador a 16°C, com a velocidade de 160 rpm. Após 15 minutos foram adicionados 250 µL de IPTG 1M, incubando a 3 horas (se a 37°C) ou durante a noite (16°C). Após esse período, as culturas foram centrifugadas a 5.000 rpm por 5 minutos a 4°C e as células ressuspensas em 20 mL de tampão de sonicação (50 mM de Tris-HCl pH 7,5 e 50 mM de NaCl), sendo lisadas em um sonicador por 20 ciclos de 15 segundos com 15 segundos de intervalo, mantendo o tubo sempre no gelo. O extrato obtido foi centrifugado por 30 minutos, a 30.000g e 4°C. O sobrenadante foi carregado em coluna HiTrap Chelating-Ni²⁺ (GE Healthcare) de 1 ml acoplada a uma bomba peristáltica, previamente carregada com NiCl₂ como descrito pelo fabricante. Feito isso, a eluição foi feita com gradiente descontínuo de imidazol, sendo o tampão composto por 50 mM Tris-HCl pH 7,5, 50 mM de NaCl e concentrações crescentes de imidazol (10, 50, 100, 300 e 500mM). Foi utilizado na eluição 5 mL para cada um dos tampões. Durante a eluição, foram coletadas frações de 1 mL. Essas frações foram analisadas em eletroforese SDS-PAGE (volume aplicado: 10 µL). As frações contendo proteína recombinante com maior grau de pureza foram unidas, submetidas à diálise (tampão de diálise: 50 mM Tris-HCl pH 7,5; 50 mM de NaCl e 20% v/v de glicerol), aliquotadas e armazenadas a -70°C.

4.6.2 SDS-PAGE - Eletroforese de proteínas sob condição desnaturante

Eletroforese de proteínas sob condições desnaturantes (contendo SDS) foi realizada em gel de poliacrilamida como descrito por LAEMMLI (1970). A concentração do gel utilizada foi a de 10%. As soluções utilizadas estão mostradas na tabela 7.

Tabela 7: soluções utilizadas na eletroforese de proteínas/SDS-PAGE.

Solução	Composição	Armazenagem
Acrilamida	30% Acrilamida (acrilamida:bis-acrilamida 37, 5:1)	4°C
Solução 1	1,5 M Tris-HCl pH 8,8; 0,3% SDS	4°C
Solução 2	0,5 M Tris-HCl pH 6,8; 0,4% SDS	4°C
10% PA	Persulfato de amônio 10% v/v em água	4°C
Tampão de Amostra	2% SDS; 10% glicerol; 0,01% azul de bromofenol; 0,0625 M Tris-HCl pH 6,8; 5% β -mercaptoetanol	4°C
Tampão de Corrida (10X)	30 g Tris-base, 140 g glicina, 10 g SDS por litro	temperatura ambiente
Solução Corante - Comassie Blue	0,5 g Comassie Blue R-250, 454,5 mL de metanol, 90,9 mL de ácido acético glacial, 454,5 mL de água destilada	temperatura ambiente
Solução Descorante	454,5 mL de metanol, 90,9 mL de ácido acético glacial, 454,5 mL de água destilada	temperatura ambiente

Os volumes das soluções utilizadas nos géis de empilhamento e de separação estão mostrados na tabela 8.

Tabela 8: volumes utilizados no preparo dos géis de empilhamento e separação.

GEL DE EMPILHAMENTO (A 4,5%)	
Acrilamida	0,75 mL
Solução 2	1,25 mL
Água	3,0 mL
10% PA	50 μ L
TEMED	5 μ L
GEL DE SEPARAÇÃO (A 10%)	
Acrilamida	3,25 mL
Solução 1	2,5 mL
Água	4,25 mL
10% PA	100 μ L
TEMED	10 μ L

As eletroforeses foram realizadas em sistema vertical seguindo as recomendações do fabricante das cubas de eletroforese (sistema *MiniProtean*, Biorad). As corridas foram realizadas a 160V, em tampão de corrida 1X por 1 hora. Após a corrida, o gel foi corado com solução de Comassie Blue por uma hora,

seguido de descoloração com solução descorante até permitir a visualização das bandas protéicas.

4.6.3 Dosagem de proteína

As concentrações de proteínas obtidas foram determinadas utilizando o método de BRADFORD (1976), onde se usou albumina de soro bovino para se obter a curva padrão.

4.6.4 Identificação de proteínas por espectrometria de massas

A identificação de proteínas por espectrometria de massas foi realizada de acordo com SHEVCHENKO e colaboradores (1996) com modificações. As bandas proteicas de interesse foram coletadas de um gel SDS-PAGE e descoradas em tubos de plástico por sucessivas lavagens com solução de bicarbonato de amônio a 25 mM pH 8,0 e 50% de acetonitrila. Posteriormente, os fragmentos foram desidratados com acetonitrila 100% e submetidos a digestão enzimática com tripsina (grau de sequenciamento de proteínas, Promega). Os fragmentos foram reidratados em 15 µL de solução contendo tripsina 15 ng/µL em tampão bicarbonato de amônio 40 mM pH 8 e acetonitrila 10% durante 30 minutos a 4°C (penetração da tripsina no gel) e, em seguida, este sistema foi incubado durante 16 horas a 37 °C (digestão propriamente dita). Após a incubação, 1 µL do sistema foi misturado com 1 µL de uma solução de acetonitrila 50% e ácido trifluoracético 0,1% saturada com a matriz HCCA (ácido α -ciano-4-hidroxicinamico). As amostras foram aplicadas em uma placa de MALDI até cristalizarem (secagem a temperatura ambiente). Os espectros foram adquiridos usando um espectrômetro MALDI-TOF-TOF Autoflex II (Bruker Daltonics). A aquisição das massas foi realizada no modo positivo utilizando voltagem de 20 kV. A lista de picos foi criada usando o software FlexAnalysis 3.0 (Bruker Daltonics). A identificação das proteínas foi realizada usando o software Mascot 2.2 e o banco de dados do GenBank.

4.7 DETERMINAÇÃO ATIVIDADE DE GLICOSIL-HIDROLASE PELO MÉTODO DO 3,5-DINITROSALICILATO

O teste do 3,5-dinitrossalicilato (DNS) (MILLER, 1959; WOOD e BHAT, 1988) permite a detecção de açúcares redutores em uma amostra. Podemos definir como açúcar redutor aquele carboidrato que pode formar grupo aldeído ou cetona quando em solução, por isomerismo. O método (figura 5) baseia-se na oxidação do açúcar redutor e redução em meio alcalino do 3,5-dinitrossalicilato (de coloração amarela). O produto formado é estável, possui coloração alaranjada e apresenta máxima absorção no comprimento de onda de 535-540 nm.

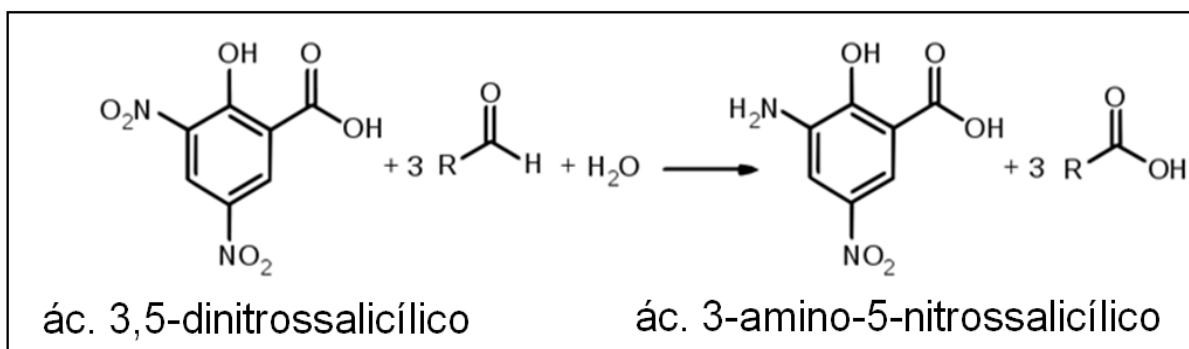


Figura 5: O ensaio de dinitrossalicilato fundamenta-se na redução do DNS, em meio alcalino, para ác. 3-amino-5-nitrossalicílico.

Como há a necessidade do uso de uma curva padrão (figura 6) utilizando glucose ou xilose, o resultado do método é expresso em açúcares redutores equivalentes a glucose ou a xilose (respectivamente), ou seja, durante a hidrólise de um polímero, por exemplo, avalia-se o surgimento de pontas redutoras e não a quantificação apenas de monômeros formados.

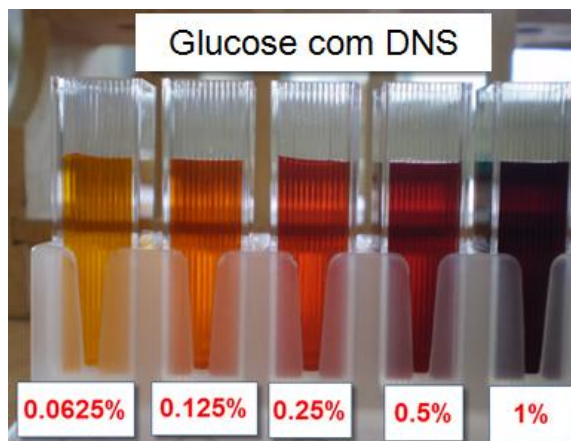


Figura 6: exemplo de curva padrão para o teste de DNS. Fonte: KOK, 2011.

Esse método é muito utilizado atualmente na determinação da atividade de celulasas e hemicelulasas, sendo recomendado pela IUPAC. Porém, uma das desvantagens é a sua baixa sensibilidade (limite de detecção: 100 μ g de glucose em 500 μ L de solução). Para avaliar baixos teores de açúcar redutor, há a necessidade de suplementação de glucose na amostra analisada.

O teste de atividade glicolítica foi feito utilizando 250 μ L de substrato (a 1% em tampão correspondente) em 25 a 250 μ L de enzima (purificada ou em extrato). Após o tempo de reação, adicionou-se 500 μ L do reativo DNS (1 g do ácido 3,5-dinitrosalicílico, 22,3 g de tartarato de potássio e sódio e 1,6 g de NaOH em 100 mL de água destilada) fez-se a homogeneização e a mistura foi mantida por 5 minutos a 100°C. Depois de resfriado em banho de gelo, foi acrescentado às amostras 4 mL de água destilada e lidas a 540 nm. Depois de descontado as absorbâncias do substrato e de enzima adicionada (feito em tubos separados), a concentração de açúcares redutores equivalentes foi obtida pela projeção sobre curva padrão, construída empregando-se glucose ou xilose.

5 RESULTADOS E DISCUSSÃO

5.1 GLICOSIL-HIDROLASES DE *AZOSPIRILLUM BRASILENSE* SP245

A estirpe Sp245 de *A. brasilense*, cujo genoma já está publicado, foi utilizada neste trabalho como fonte de glicosil-hidrolases. WISNIEWSKI-DYÉ e colaboradores (2011), após analisarem o genoma dessa bactéria, identificaram 15 genes que possuem regiões conservadas para celulasas e hemicelulasas. Destes, 4 genes não apresentam domínios completos para glicosil-hidrolases e um gene possui similaridade com uma β -mananase. Assim foram escolhidos apenas 10 genes para serem estudados neste projeto. Estas enzimas foram classificadas através da comparação de sequências com o banco de dados GenBank (<http://www.ncbi.nlm.nih.gov>) através do programa BLASTx (ALTSCHUL *et al.*, 1997) (tabela 9). Para a obtenção de mais dados sobre os genes, foram utilizados também alguns programas presentes no site ExPASy (<http://www.expasy.org/>): *ProtParam* (obtenção teórica de parâmetros físicos e químicos - GASTEIGER *et al.*, 2005), *ScanProsite* (busca por domínios conservados - SIGRIST *et al.*, 2002) e *PSORT* (predição de localização protéica – NAKAI e KANEHISA, 1991).

Tabela 9: Genes de prováveis glicosil-hidrolases de *A. brasilense* Sp245 e características de suas respectivas proteínas.

Nº	Denominação	ID ¹	Família ¹	Atividade ¹	Nº (aa) ²	pI ²	MW(kDa) ^{2*}
1	XilanaseA	AZOBR_p1110166	GH10	Xilanase	419	7,7	49,1
2	XilanaseB	AZOBR_p310014	GH10	Xilanase	378	8,4	43,0
3	CelulaseA	AZOBR_p440082	GH12	Endoglucanase	272	4,1	31,1
4	CelulaseB	AZOBR_p140078	GH1	β -glucosidase	482	6,3	55,0
5	CelulaseC	AZOBR_p280124	GH1	β -glucosidase	444	5,2	51,6
6	CelulaseD	AZOBR_p280134	GH1	β -glucosidase	457	5,7	52,0
7	CelulaseE	AZOBR_p280136	GH1	β -glucosidase	376	5,4	44,1
8	CelulaseF	AZOBR_150049	GH5	Endoglucanase	441	4,7	50,0
9	CelulaseG	AZOBR_p470008	GH5	Endoglucanase	689	4,46	75,0
10	CelulaseH	AZOBR_p1110164	GH5	Endoglucanase	825	4,8	91,2

Nº	Domínios conservados ^{1,4}	Tipo ³
1	Família GH12	Citoplasmática
2	sinal de secreção (TAT <i>signal profile</i>) e família GH1	Extracelular
3	sinal de secreção (TAT <i>signal profile</i>) e família GH10	Extracelular**
4	família GH10	Extracelular
5	família GH1	Citoplasmática
6	família GH1	Citoplasmática
7	família GH1	Citoplasmática
8	família GH5	Citoplasmática
9	família GH5	Citoplasmática
10	família GH5 e domínio CalX- β ***	Citoplasmática

Legenda: 1 – Obtido com BlastX, 2 – dados do ProtParam, 3 – dados do PSORT, 4 – dados do ScanProsite, * - massa da proteína com cauda de histidina,

** - predito citoplasmática pelo PSORT, mas apresenta sinal TAT de exportação, *** - domínio envolvido na regulação intracelular de cálcio.

Como visto na tabela 9, os genes 2 (XilanaseB) e 3 (CelulaseA) apresentam sequência sinal de secreção TAT dependentes (do inglês *twin-arginine translocation*). Esta sequência sinal é rica em aminoácidos básicos (S-R-R-X-F-L-K) e permite o transporte da proteína alvo já enovelada do ambiente citoplasmático para o meio extracelular (DE BUCK *et al.*, 2008). Como estes genes foram clonados em vetor pET28a, o programa PSORT (NAKAI e KANEHISA, 1991) indicou que as proteínas recombinantes se tornaram citoplasmáticas após a inserção de cauda de histidina em suas regiões N-terminais.

A partir das sequências nucleotídicas, foram sintetizados iniciadores específicos (tabela 4) com o intuito de se obter os genes utilizando a técnica de PCR e o DNA genômico de *A. brasilense* Sp245 como molde. Além disso, as sequências dos iniciadores foram modificadas para se criarem sítios de clivagem para enzimas de restrição, facilitando a posterior clonagem em vetores de expressão.

Não foi possível a criação de sítios para *BamHI* nos iniciadores para os genes de número 2, 3 e 10, devido a sítios de clivagem presentes no interior das sequências nucleotídicas. Assim, optou-se pela utilização de sítio da enzima *XhoI* nesses casos.

5.2 CLONAGEM DOS GENES DE *AZOSPIRILLUM BRASILENSE* SP245

Após a seleção dos genes de *A. brasilense* Sp245, a primeira etapa realizada foi a amplificação das sequências através de técnica de PCR, utilizando DNA genômico como molde. Seguindo o protocolo descrito na seção Material e Métodos, não se detectou qualquer amplificação. Assim, algumas tentativas foram feitas modificando as seguintes condições de reação (feitas em conjunto ou separadamente):

- a) Variação da temperatura de anelamento em 5 ou 10 °C (para mais ou para menos) em relação à teórica;
- b) Uso de tampão da Taq DNA polimerase contendo $(\text{NH}_4)_2\text{SO}_4$ (0,2M) ao invés de KCl (50mM);
- c) Adição de 1, 5 e 10 % (v/v) de DMSO (dimetilsulfóxido) à reação;
- d) Variação da concentração de DNA molde na reação: 1, 2,5, 5, 10, 200 e 250 ng;

- e) Diminuição gradativa da temperatura de anelamento durante os ciclos da PCR (técnica *Touchdown* PCR).

Ainda assim não se detectava o aparecimento de amplificados. Essa dificuldade pode estar relacionada ao elevado conteúdo de bases G e C (guanina e citosina) presente no genoma de *A. brasilense* Sp245, que atinge 68,5% (WISNIEWSKI-DYÉ *et al.*, 2011). Regiões ricas em G+C tendem a formar estruturas secundárias, principalmente grampos, que impedem a atividade da DNA polimerase na PCR (HAQQI *et al.*, 2002). Um dos compostos que podem desfazer tais estruturas é o DMSO, que diminui as interações existentes entre as bases nitrogenadas, facilitando a separação das fitas de DNA (FRACKMAN *et al.*, 1998), porém seu uso neste trabalho não gerou resultados positivos.

A betaína é outro composto que diminui a formação de estrutura secundária de ácido nucleico de fita simples. Quando se adicionou betaína (na concentração de 1 M) na reação de PCR, obteve-se resultado positivo para os genes xilanaseA, xilanaseB, celulaseB, celulaseF e celulaseG (figura 7).

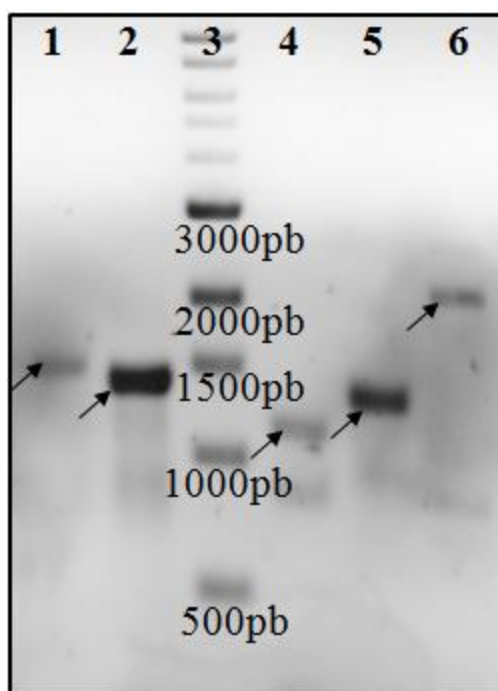


Figura 7: reação de PCR com 1 M de betaína. Gel de agarose 1% evidenciando os amplificados obtidos. 1: amplificado referente ao gene celulaseB (1475 pb). 2: amplificado referente ao gene xilanaseA (1379 pb). 3: marcador de tamanho molecular. 4: amplificado referente ao gene xilanaseB (1168 pb). 5: amplificado referente ao gene celulaseF (1354 pb). 6: amplificado referente ao gene celulaseG (2112 pb).

Acredita-se que a betaína diminui a força de interação entre bases GC, o que levaria a desestruturação de grampos no DNA, permitindo o acesso da DNA polimerase. As interações entre bases GC diminuiriam pelo aumento de hidratação da fita causada pela betaína, igualando as forças de pareamento entre bases GC e AT (MAMMEDOV *et al*, 2008).

Os produtos de PCR obtidos foram ligados ao vetor pTZ57R/T, através de sistema de ligação do tipo TA. Feito isso, os produtos de ligação foram transformados em *E. coli* DH10B e plaqueados em meio LA contendo ampicilina (100 ug/mL) e X-gal (30 ug/mL). As colônias brancas foram submetidas à técnica de PCR de colônia utilizando os iniciadores M13 universal e reverso para confirmar a presença do gene de interesse. Como controle negativo, foram coletadas colônias com coloração azul (vetor vazio, ou seja, gene *lacZ* íntegro).

Como não houve sucesso na clonagem do gene XilanaseB, ela não será discutida no decorrer do trabalho.

5.2.1 Clonagem do gene celulaseB de *A. brasilense* Sp245

O gene celulaseB (gene número 4 na tabela 9, página 49) codifica para uma proteína da classe GH1 das glicosil-hidrolases. Uma vez que o tamanho do amplificado contendo este gene é 1475 pb, clones que apresentarem banda em torno de 1630 pb (gene mais sequência do vetor) após a PCR de colônia serão aqueles que possuem o plasmídeo pTZ57R/T contendo o gene celulaseB. Esse aumento de tamanho se deve ao uso de iniciadores específicos para esse vetor (iniciadores M13), que se anelam a uma região mais afastada daquela onde o produto de PCR foi ligado. Assim, há o incremento de 153 pb no tamanho da banda esperada.

Após a ligação do vetor com o produto de PCR do gene celulaseB e transformação em *E. coli* DH10B, houve o aparecimento de 8 colônias sem coloração azul. Esses clones foram submetidos ao teste de PCR de colônia e os resultados estão mostrados a seguir (figura 8).

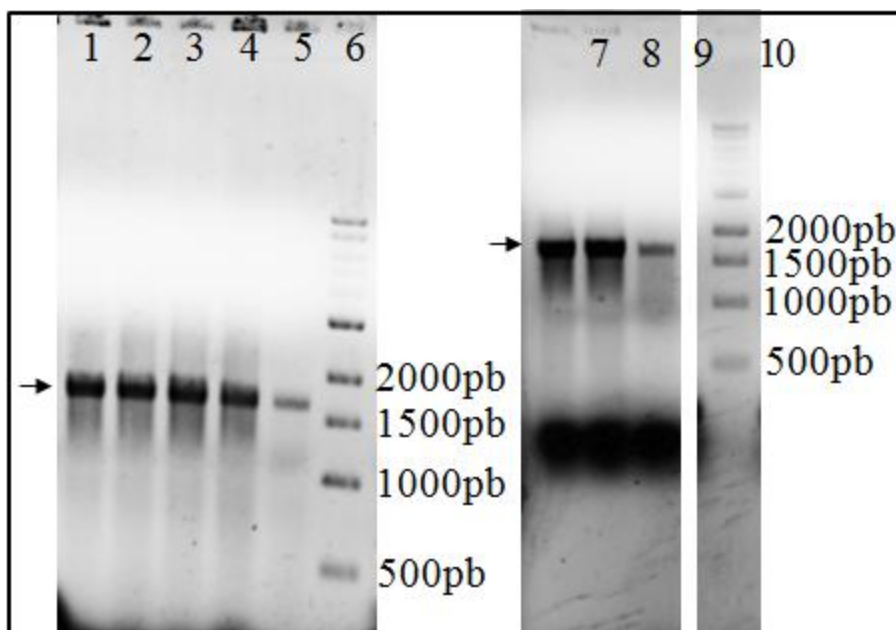


Figura 8: confirmação dos clones transformantes após ligação pTZ57R/T + gene celulaseB. Gel de agarose 1% mostrando o resultado da PCR de colônia. Poços 1 a 5 e 7 a 9: 8 colônias brancas. Poços 6 e 10: marcador de tamanho molecular. As setas mostram a banda esperada que indica a presença do gene.

Esse experimento demonstrou que as 8 colônias coletadas geraram uma banda próxima a 1630 pb após PCR, como esperado. Para analisar o tamanho do inserto, foi realizada a extração plasmidial e posterior digestão com as enzimas *EcoRI* e *HindIII* com liberação do inserto no vetor pTZ57R/T. A figura 9 mostra o padrão de restrição de 3 clones.

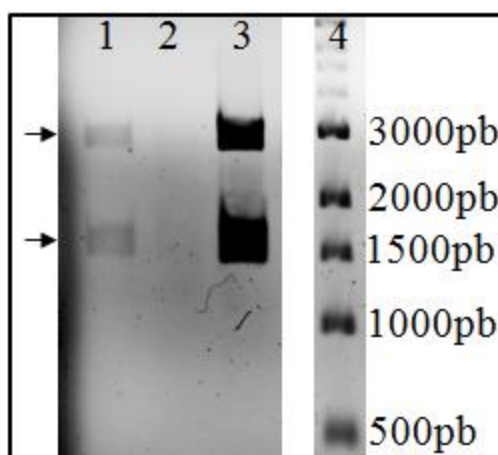


Figura 9: Gel de agarose 1% mostrando resultado da restrição com as enzimas *EcoRI* e *HindIII* de 3 plasmídeos derivados do vetor pTZ57R/T + inserto celulaseB. Poços 1 a 3: plasmídeos isolados de 3 colônias transformantes. Poço 4: marcador de tamanho molecular. As setas indicam as duas bandas esperadas (vetor com ~3000 pb e inserto com 1500 pb).

Analisando o resultado da restrição, 2 clones provavelmente apresentam o fragmento do gene celulaseB, uma vez que apresentam um fragmento *EcoRI/HindIII*

de aproximadamente 1500 pb (inserto) e outro em torno de 3000 pb (vetor). O poço 2 apresentou bandas muito fracas em relação aos poços vizinhos, mas ainda assim aparenta ter as duas bandas características. Um dos clones foi estocado em glicerol 50% a -20°C e seu plasmídeo foi denominado pMS_zCelB, cujo inserto foi sequenciado. A comparação feita entre a sequência do gene CelulaseB clonado com a sequência genômica de *A. brasilense* Sp245 está mostrado na figura 10.

Alinhamento do gene celulaseB (CelB) presente no genoma (G) e o clonado (C):

```

CelB G AGGAATTTCGGTATGCGGCGGCGGAGTTTCCTGAAGGCGGCGGCGGCCACCGGGCTGGCGG 60
CelB C AAGAATTCCATATGCGGCGGCGGAGTTTCCTGAAGGCGGCGGCGGCCACCGGGCTGGCGG 60
      * .***** .*****

CelB G CCTCGGCCGGGGCCTTTCTGACCACGCGCCGTGCCGCGGTGGCCCAGCCGTCGGAAGTGC 120
CelB C CCTCGGCCGGGGCCTTTCTGACCACGCGCCGTGCCGCGGTGGCCCAGCCGTCGGAAGTGC 120
      *****

CelB G TGGCCGCCTTTCCCGACAAGTTCGTCTGGGGTGCCTCGACCTCCTCCTACCAGATCGAGG 180
CelB C TGGCCGCCTTTCCCGACAAGTTCGTCTGGGGTGCCTCGACCTCCTCCTACCAGATCGAGG 180
      *****

CelB G GGGCGGTCACCGCCGGCGGGCGCGGCCCCAGCGTGTGGGACACCTTCAGCCATTCCTTCG 240
CelB C GGGCGGTCACCGCCGGCGGGCGCGGCCCCAGCGTGTGGGACACCTTCAGCCATTCCTTCG 240
      *****

CelB G GCAAGGTGGCGAACGGCGACACCGGCGACGTGGCCTGCGACCATTACAACCGCTACGCCG 300
CelB C GCAAGGTGGCGAACGGCGACACCGGCGACGTGGCCTGCGACCATTACAACCGCTACGCCG 300
      *****

CelB G AGGATGTGGACCTGATGGCGAAGGCGGGGATGAACGCCTATCGCTTCTCCGTCGCCTGGC 360
CelB C AGGATGTGGACCTGATGGCGAAGGCGGGGATGAACGCCTATCGCTTCTCCGTCGCCTGGC 360
      *****

CelB G CGCGCGTCCAGCCGACCGGCACCGGCCCGGCCAACGCCGAGGGGCTGGACTTCTACGACC 420
CelB C CGCGCGTCCAGCCGACCGGCACCGGCCCGGCCAACGCCGAGGGGCTGGACTTCTACGACC 420
      *****

CelB G GGCCGACCGACGCGCTGCTGGCCAAGGGCATCGCGCCCTGGCCCTGCCTGTACCATTGGG 480
CelB C GGCTGACCGACGCGCTGCTGGCCAAGGGCATCGCGCCCTGGCCCTGCCTGTACCATTGGG 480
      *** *****

CelB Gen ACCTGCCCCAGGCGCTCCAGGACCGCGGCGGCTGGACCAATCGGGACATCGCCGGCTGGT
540
CelB Clo ACCTGCCCCAGGCGCTCCAGGACCGCGGCGGCTGGACCAATCGGGACATCGCCGGCTGGT
540
      *****

```

(continua)

CelB G	TCACCGACTACGCCCAGCTTGTGCGCCGCGCGGATCGGCGACCGAGCCAAGCACTGGACCA	600
CelB C	TCACCGACTACGCCCAGCTTGTGCGCCGCGCGGATCGGCGACCGAGCCAAGCACTGGACCA	600

CelB G	TGCTGAACGAGCCGTCCGTCCACGCCATCTTCGGCCACGGGCTGGGCGGGCACGCGCCGG	660
CelB C	TGCTGAACGAGCCGTCCGTCCACGCCATCTTCGGCCACGGGCTGGGCGGGCACGCGCCGG	660

CelB G	GGATGACCGGCAAGGGCAACTATTTCAAGGCGATCCACCACCAGAACCTCGCCCAGGGCA	720
CelB C	GGATGACCGGCAAGGGCAACTATTTCAAGGCGATCCACCACCAGAACCTCGCCCAGGGCA	720

CelB G	TGGCGCTGAAGGCTCTGCGCGCCGCGGGCGGGGCCAAGGGCTGGCAGCTCGGCACCGTGC	780
CelB C	TGGCGCTGAAGGCTCTGCGCGCCGCGGGCGGGGCCAAGGGCTGGCAGCTCGGCACCGTGC	780

CelB G	TGTCGCTCCAGCCGGTCTGGCCGGTTCGGCGGGCTGGACGCCAATTACGCGGCCTCGCTGA	840
CelB C	TGTCGCTCCAGCCGGTCTGGCCGGTTCGGCGGGCTGGACGCCAATTACGCGGCCTCGCTGA	840

CelB G	TGTGGGACGCGGTGTGGAACCGCGCCTGCCTCGACCCGCTGCTGCGTGGGGAGTATCCGG	900
CelB C	TGTGGGACGCGGTGTGGAACCGCGCCTGCCTCGACCCGCTGCTGCGTGGGGAGTATCCGG	900

CelB G	AACTGCTGCGCGATGGCTTCGCCCCCTGGTCAAGGCCGGCGACCTGGAGGCGATCCGCC	960
CelB C	AACTGCTGCGCGATGGCTTCGCCCCCTGGTCAAGGCCGGCGACCTGGAGGCGATCCGCC	960

CelB G	AGCCCATCGACTTCCTGGGCATCAACTACTACAGCCGGATGCACCAGCAGCCGGACCCGG	1020
CelB C	AGCCCATCGACTTCCTGGGCATCAACTACTACAGCCGGATGCACCAGCAGCCGGACCCGG	1020

CelB G	CGGGCCTGTTTCGGCACCGGCTACGGCTCCCCGCCGGAAGGCACGCCGACGACCGGGATGG	1080
CelB C	CGGGCCTGTTTCGGCACCGGCTACGGCTCCCCGCCGGAAGGCACGCCGACGACCGGGATGG	1080

CelB G	GCTGGCCGGTGGAGCCCCGACGGCATCGCCGAGATCCTGATCGAGCTGAAGCAGGAGTACG	1140
CelB C	GCTGGCCGGTGGAGCCCCGACGGCATCGCCGAGATCCTGATCGAGCTGAAGCAGGAGTACG	1140

CelB G	GCAACCCCCCGGTCTACGTGATGGAGAACGGTGCCGCCTATCCCGAGCAGACCGGGCCGA	1200
CelB C	GCAACCCCCCGGTCTACGTGATGGAGAACGGTGCCGCCTATCCCGAGCAGACCGGGCCGA	1200

CelB G	AGGGCTTCGTCCAGGACAACGACCGCATCAGCTACCTGCGCCGCCACATCCTGGCCGGTTC	1260
CelB C	AGGGCTTCGTCCAGGACAACGACCGCATCAGCTACCTGCGCCGCCACATCCTGGCCGGTTC	1260

CelB G	ATCAGGCGTTGGAGGAGGGGGTCGACCTGCGGGGCTGGTTTCGTGTGGAGCCTGCTGGAC	1320
CelB C	ATCAGGCGTTGGAGGAGGGGGTCGACCTGCGGGGCTGGTTTCGTGTGGAGCCTGCTGGAC	1320

CelB G	ACTTCGAGTGGGCCGAGGGTTACCAGCGCCGCTTCGGCCTGATCGAGGTGGACCGCCAGA	1380
CelB C	ACTTCGAGTGGGCCGAGGGTTACCAGCGCCGCTTCGGCCTGATCGAGGTGGACCGCCAGA	1380

(continua)

CelB	G	CCCTGGAGCGCCGTCCCAAGGCCAGCTACCACTGGTACGCCGACGTGATCCGCAAGAAGG	1440
CelB	C	CCCTGGAGCGCCGTCCCAAGGCCAGCTACCACTGGTACGCCGACGTGATCCGCAAGAAGG	1440

CelB	G	GCGTTCGACCTCGGCTTGAGAGGCTCGGTTCAAG	1475
CelB	C	GCGTTCGACCTCGGCTTGAGAGGCTCGGATCCAG	1475
*****:*. **			

Figura 10: Comparação de sequências do gene CelulaseB (CelB) presente no genoma de *A. brasilense* Sp245 com o inserto clonado. Programa ClustalW2 (LARKIN *et al.*, 2007). * para bases iguais, : para similares e . para pouco similares. As bases sublinhadas indicam a localização dos iniciadores utilizados na clonagem. Bases marcadas em cinza evidenciam o início e o fim do gene.

A comparação de sequências do inserto com o genoma de *A. brasilense* mostrou duas mutações (além das 5 bases modificadas nos iniciadores para obtenção de sítios de restrição), causando as modificações L138P (resíduo de leucina 138 substituído por resíduo de prolina) e N437D (resíduo de asparagina 437 para resíduo de aspartato; resíduo polar para resíduo polar com carga negativa).

O próximo passo realizado foi a passagem do inserto para o vetor de expressão pET28a. Assim, o plasmídeo pMS_zCelB e o vetor pET28a foram digeridos com *NdeI* e *BamHI*. O inserto de aproximadamente 1,5 Kb foi separado por eletroforese em agarose, extraído e purificado do gel, e ligado ao vetor digerido com T4 DNA ligase. Depois, o produto de ligação foi transformado em *E. coli* DH10B.

A figura 11 mostra o produto de PCR das 3 colônias transformantes, utilizando iniciadores para o promotor e terminador T7 (sítios presentes no pET28a; o uso deste iniciadores aumenta em aproximadamente 300 pb a banda esperada).

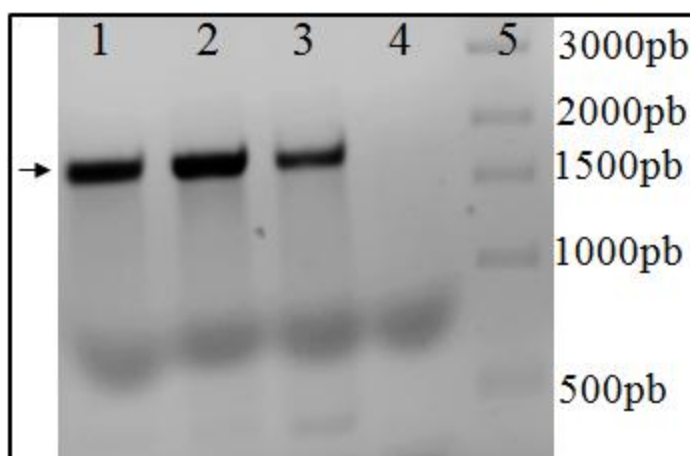


Figura 11: Confirmação da presença do gene CelulaseB no vetor pET28a – gel de agarose 1%. PCR de colônia realizado com iniciadores T7. Bandas em torno de 1775 pb apontam resultado positivo (indicadas pela seta). Poços 1 a 3: clones transformantes. Poço 4: controle negativo (clone com vetor vazio). Poço 5: marcador de tamanho molecular.

O resultado obtido foi positivo para os 3 clones. Assim, um dos clones contendo o gene CelulaseB (com mutações) no vetor pET28a, denominado pMS_CelB, foi selecionado e estocado em glicerol 50% a -20°C.

5.2.2 Clonagem do gene XilanaseA de *A. brasilense* Sp245

O gene XilanaseA (gene número 1 na tabela 9, p. 42) de *A. brasilense* estirpe Sp245 codifica para uma proteína pertencente à classe GH10 das glicosil-hidrolases. Este gene foi amplificado por PCR gerando um fragmento de 1379 pb.

Após a ligação do produto de PCR de XilanaseA no vetor pTZ57R/T e posterior transformação em *E. coli* DH10B, houve o aparecimento de 15 colônias brancas em LA contendo ampicilina e Xgal (Material e Métodos, item 4.2). Sete dessas colônias foram submetidas ao teste de PCR de colônia, cujo resultado permitiu inferir que nenhuma delas continha o plasmídeo com o inserto esperado. Dessa forma, as outras 8 colônias foram submetidas ao mesmo teste, e o resultado está mostrado na figura 12.

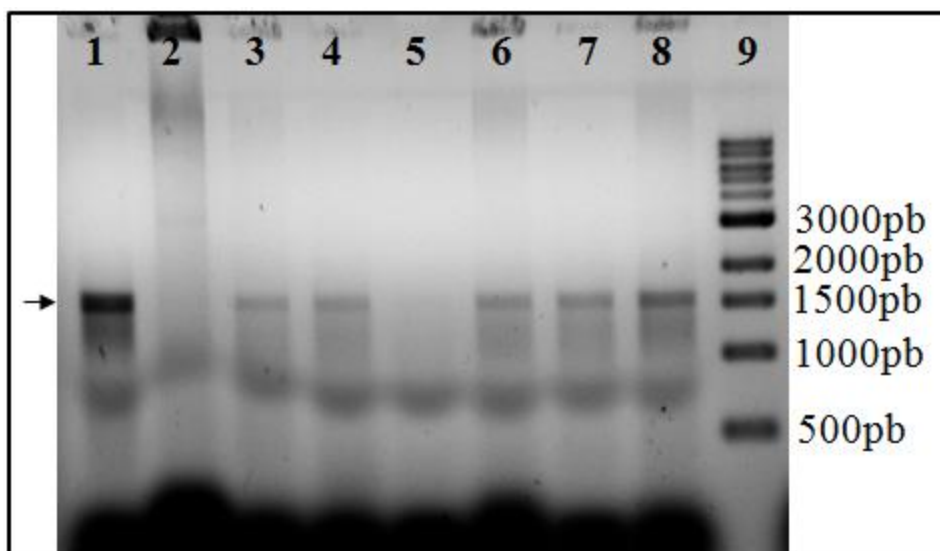


Figura 12: Identificação de colônias transformantes contendo o vetor pTZ57R/T com gene XilanaseA. Eletroforese em gel de agarose 1% mostrando o resultado da PCR de colônia utilizando iniciadores M13 universal e reverso. Poço 1 a 8: colônias transformantes. Poço 9: marcador de tamanho molecular. A seta indica a banda esperada (~1500 pb) para plasmídios contendo o gene.

Das 8 colônias, 6 apresentaram uma banda próxima a 1500 pb. Para confirmar o tamanho do inserto, realizou-se a extração e digestão do DNA plasmidial

com as enzimas *EcoRI* e *HindIII*. Como o gene XilanaseA apresenta sítio de clivagem para *EcoRI*, esperava-se o aparecimento de uma banda de 705 pb e outra de 825 pb, além da banda referente ao vetor aberto (próximo de 3000 pb). Foram escolhidos 3 dos 6 clones para seguir com as etapas de extração plasmidial e digestão enzimática (figura 13).

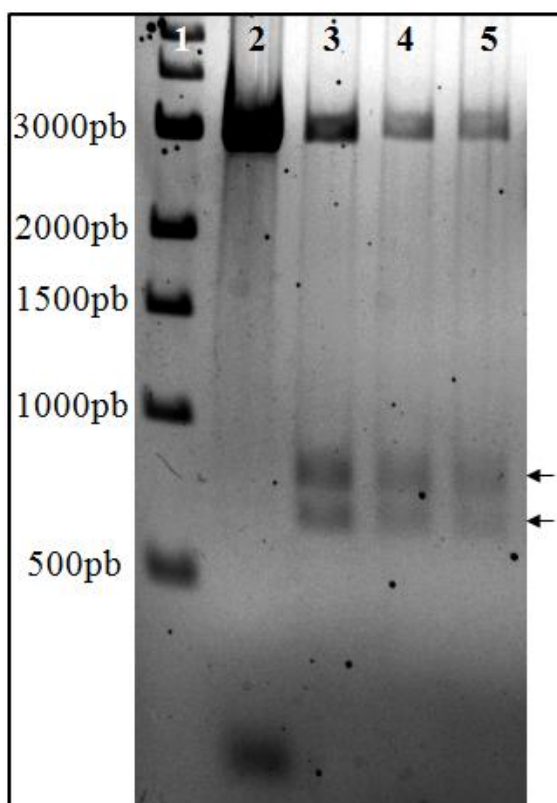


Figura 13: Análise de restrição de 3 plasmídeos pTZ57R/T+XilanaseA, utilizando as enzimas *EcoRI* e *HindIII* – gel de agarose 1%. Poço 1: marcador de tamanho molecular. Poço 2: colônia com vetor pTZ57R/T vazio (controle negativo). Poços 3 a 5: plasmídeos de 3 clones transformantes após ligação do vetor pTZ57R/T com gene XilanaseA. As setas mostram a localização das bandas esperadas (825 e 705 pb).

O resultado mostra que os 3 clones testados podem conter o fragmento do gene XilanaseA. Os insertos dos 3 clones foram extraídos e ligados ao vetor de expressão pET28a, porém apenas um deles gerou clones transformantes. Este inserto, cujo plasmídeo de origem passou a se chamar pMS_zXilA, foi sequenciado e comparado com a sequência do gene XilanaseA no genoma de *A. brasilense* Sp245 (figura 14).

Alinhamento do gene XilanaseA (XilA) presente no genoma (G) e o clonado (C):

```

Xila G   ACGCCCATGGCCCAGACCTCGCCAAAAGCCACCTCGCCGAAAGCCCTCTCCTTCACGCGC 60
Xila C   ACGCATATGGCCCAGACCTCGCCAAAAGCCACCTCGCCGAAAGCCCTCTCCTTCACGCGC 60
        ****.*****

Xila G   CGAGACGCCTTGCGCGCCGCCGCGAGCCGCCGCGGTGCTGGCAAGCTGCAAGGCCGCCGCC 120
Xila C   CGAGACGCCTTGCGCGCCGCCGCGAGCCGCCGCGGTGCTGGCAAGCTGCAAGGCCGCCGCC 120
        *****

Xila G   GGTCCCGCGTTCGCCGAGGCGGCGTCCGTTGCCGCCGTGCCCAACCCACCCCTGCGCGAC 180
Xila C   GGTCCCGTCGTCGCCGAGGCGGCGTCCGTTGCCGCCGTGCCCAACCCACCCCGCGCGAC 180
        *****

Xila G   GCCTGCCGGGCGCCGGGATTCTTCACGGCGCCGCCGCGACCACTGATCGAGCCGGAA 240
Xila C   GCCTGCCGGGCGCCGGGATTCTTCACGGCGCCGCCGCGACCACTGATCGAGCCGGAA 240
        *****

Xila G   GACCCCATGCTCGACACGCTGATGGCGCGGGAATGCGACGTGGTGA CTCCGAGAACGGC 300
Xila C   GACCCCATGCTCGACACGCTGATGGCGCGGGAATGCGACGTGGTGA CTCCGAGAACGGC 300
        *****

Xila G   GGCAAATGGGCCGTCTTCCAGCCGCGAGGAGGGCAACTTCGACTGGGGCCGCTTCGACGCC 360
Xila C   GGCAAATGGGCCGTCTTCCAGCCGCGAGGAGGGCAACTTCGACTGGGGCCGCTTCGACGCC 360
        *****

Xila G   GGGGTGGAGCTGGCCCGCCGCATCGGCGCCAAGCCCAACTGGCACTGCGCCCTGTGGCAG 420
Xila C   GGGGTGGAGCTGGCCCGCCGCATCGGCGCCAAGCCCAACTGGCACTGCGCCCTGTGGCAG 420
        *****

Xila G   CACATGGGCATGCCCCGACTACATGACGCTGCCGGCCAGCCGCCAGCGCGCGCTGGGCATC 480
Xila C   CACATGGGCATGCCCCGACTACATGACGCTGCCGGCCAGCCGCCAGCGCGCGCTGGGCATC 480
        *****

Xila G   GCGGAGAGCGCCTATTTCTCGGCCGACGGCAGCTGTCCGAGGACAATTACTGGCCGCGC 540
Xila C   GCGGAGAGCGCCTATTTCTCGGCCGACGGCAGCTGTCCGAGGACAATTACTGGCCGCGC 540
        *****

Xila G   TTCACCGCCATGGTCGCGGCGGTCAAGCAGCGCTACGGCGACAATTTCTACCGCATCGAC 600
Xila C   TTCACCGCCATGGTCGCGGCGGTCAAGCAGCGCTACGGCGACAATTTCTACCGCATCGAC 600
        *****

Xila G   GTGATGAACGAGGCCTTCTTCTGGGAGACCGAGCGCAGCCACCCCAAGGAGCAGGACCGC 660
Xila C   GTGATGAACGAGGCCTTCTTCTGGGAGACCGAGCGCAGCCACCCCAAGGAGCAGGACCGC 660
        *****

Xila G   TACGGCTTCCGCAAGGGCATGTGGTGGGTGGTGGCCGGCGGCGCCAAGGGGCCGGAATGG 720
Xila C   TACGGCTTCCGCAAGGGCATGTGGTGGGTGGTGGCCGGCGGCGCCAAGGGGCCGGAATGG 720
        *****

Xila G   CTGGACCCGTTCTTCCACCACATCCGCAAGGAATTCCTCCTCCAAGCTGGTTATCAAC 780
Xila C   CTGGACCCGTTCTTCCACCACATCCGCAAGGAATTCCTCCTCCAAGCTGGTTATCAAC 780
        *****

```

(continua)

Xila	G	GAGTTCGGGATTGAGATCGACGAGGGCTGGCAGCAGCGCAAGCGCGCCTATTTCCAGACC	840
Xila	C	GAGTTCGGGATTGAGATCGACGAGGGCTGGCAGCAGCGCAAGCGCGCCTATTTCCAGACC	840

Xila	G	TGGCTGACCGACGCGGTCAAGCGCGGCGTGCCGATCGACGGCGTCGGGCTGCAGAGCCAC	900
Xila	C	TGGCTGACCGACGCGGTCAAGCGCGGCGTGCCGATCGACGGCGTCGGGCTGCAGAGCCAC	900

Xila	G	CTGATGGCCGGCAAGCCCTACGACCGCGAGGGCATGAAGGGCTTCCTGCGGGCCATGGAC	960
Xila	C	CTGATGGCCGGCAAGCCCTACGACCGCGAGGGCATGAAGGGCTTCCTGCGGGCCATGGAC	960

Xila	G	CGGCTGGGCCTGCCCATCCACGTCACCGAGTTTCGACGTGAACGAGAAGCACCTGCCGCGC	1020
Xila	C	CGGCTGGGCCTGCCCATCCACGTCACCGAGTTTCGACGTGAACGAGAAGCACCTGCCGCGC	1020

Xila	G	TCCTGGTTCGCGGGCGGAGAAGGACCGCGCCATGGCCTTCCTGGCCGGGCAGTATCTGGGC	1080
Xila	C	TCCTGGTTCGCGGGCGGAGAAGGACCGCGCCATGGCCTTCCTGGCCGGGCAGTATCTGGGC	1080

Xila	G	GACATCGCCGGGAACGCGCGGCTGGCCGAGCTGTGCTGGTGGCACCTGCGCTCCGACCTC	1140
Xila	C	GACATCGCCGGGAACGCGCGGCTGGCCGAGCTGTGCTGGTGGCACCTGCGCTCCGACCTC	1140

Xila	G	AACTACATCTCCCGCGAATTGCCGGAGCTGAAGCCCCATCCCTCGCCCTACGACGCGGCG	1200
Xila	C	AACTACATCTCCCGCGAATTGCCGGAGCTGAAGCCCCATCCCTCGCCCTACGACGCGGCG	1200

Xila	G	TCGCGCCCGCTGCCGCTGTACGAGGCGTCGGTGCAGGCGCTGCGCGGACGCAACCGGTCTG	1260
Xila	C	TCGCGCCCGCTGCCGCTGTACGAGGCGTCGGTGCAGGCGCTGCGCGGACGCAACCGGTCTG	1260

Xila	G	GGGTGAGGCGGATCGGGGCGACTCCGGCCGGGAAGCGGCGCGGTACCCCTTCCGGCCCCCT	1320
Xila	C	GGGTGAGGCGGATCGGGGCGACTCCGGCCGGGAAGCGGCGCGGTACCCCTTCCGGCCCCCT	1320

Xila	G	CTACCCCGCCCGATCGCGCGGCCCGCTCCGCCGTCCGCGCCTGTAAGGATTTCCGACAG	1379
Xila	C	CTACCCCGCCCGATCGCGCGGCCCGCTCCGCCGTCCGCGCCTGTAAGGATTTCCGACAG	1379

Figura 14: Comparação de sequências do gene XilanaseA (XilA) presente no genoma de *A. brasilense* Sp245 com o inserto clonado. Programa ClustalW2 (LARKIN *et al.*, 2007). * para bases iguais, : para similares e . para pouco similares. As bases sublinhadas indicam a localização dos iniciadores utilizados na clonagem. Bases marcadas em cinza evidenciam o início e o fim do gene.

A comparação apontou duas mutações (além das esperadas, presentes nos iniciadores para obtenção de sítios de restrição) que são: modificações A41V (resíduo de alanina 41 substituído por valina), e L56P (resíduo de leucina 56 para resíduo de prolina).

O plasmídeo derivado do vetor pET28a clonado com o gene XilanaseA foi denominado pMS_XilA. A cultura de *E. coli* DH10B contendo este plasmídeo foi conservada em glicerol 50% a -20°C.

5.2.3 Clonagem do gene CelulaseF de *A. brasilense* Sp245

O gene CelulaseF (gene número 8 na tabela 9, p. 42) codifica uma proteína que apresenta região conservada similar àquelas presentes em glicosil-hidrolases da classe GH5, possuindo um tamanho de 1354 pb.

Após realizada a ligação do vetor com o produto de PCR contendo o gene CelulaseF e posterior transformação em *E. coli* DH10B, coletaram-se 5 colônias sem coloração azul que foram submetidas ao teste de PCR de colônia e os resultados estão mostrados abaixo (figura 15).

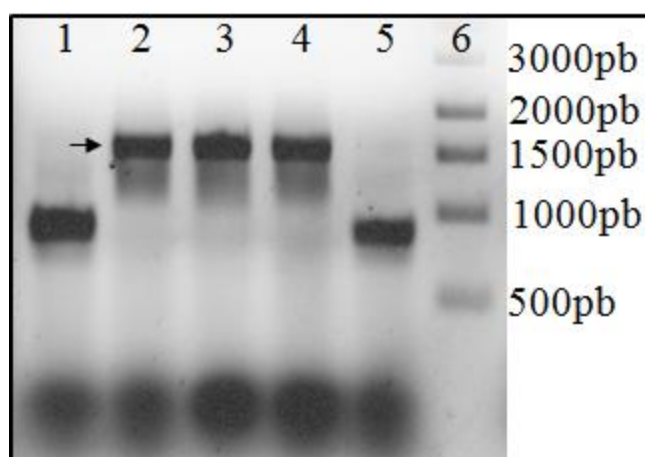


Figura 15: confirmação de clones transformantes após ligação pTZ57R/T + gene CelulaseF. Gel de agarose 1% mostrando o resultado da PCR de colônia utilizando iniciadores M13. Poço 1 a 5: clones da transformação com gene CelulaseF. Poço 6: marcador de tamanho molecular. A seta indica a possível banda relativa ao inserto CelulaseF.

Das 5 colônias analisadas, 3 apresentaram uma banda próxima a 1500 pb após a PCR. O DNA plasmidial das colônias foi extraído e analisado por digestão com as enzimas *EcoRI* e *HindIII*. Como o gene CelulaseF também apresenta sítio de clivagem para *EcoRI*, o aparecimento de uma banda próxima a 700 pb e outra de 800 pb, além da banda em torno de 3000 pb (vetor aberto), indica a presença do inserto. Apenas 2 dos 3 clones foram analisados (figura 16).

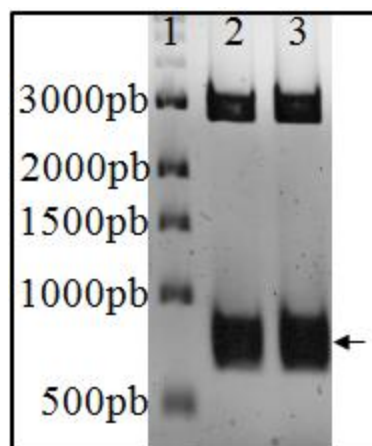


Figura 16: Análise de restrição com *EcoRI* e *HindIII* de 2 plasmídeos derivados da ligação vetor pTZ57R/T + gene CelulaseF, em gel de agarose 1%. Poço 1: marcador de tamanho molecular. Poço 2 e 3: plasmídeos de 2 clones transformantes. A seta mostra uma banda arrastada que pode representar as duas bandas esperadas (700 e 800 pb).

Como houve o aparecimento de apenas uma banda na eletroforese, os insertos foram sequenciados a fim de confirmar a presença do gene CelulaseF. O resultado mostrou que o gene está presente nos dois plasmídeos analisados. Posteriormente, foi feita a comparação entre a sequência do genoma e as dos dois plasmídeos (figuras 17 e 19).

Alinhamento do gene CelulaseF (CelF) presente no genoma (G) e o clonado (C):			
CelF G	GGCTGAGCACAATGGGCATCGCCAAGGGCTTCTGTCACACGGACGGCAACCAGATCGTCG	60	
CelF C	GGCTGAGACATATGGGCATCGCCAAGGGCTTCTGTCACACGGACGGCAACCAGATCGTCG	60	
	***** ...:*****		
CelF G	ATTTCGTCGGGTCAGAACGTCAAGCTGACGGGCGTCAACTGGTTTCGGCGGCGAGGGCTTCG	120	
CelF C	ATTTCGTCGGGTCAGAACGTCAAGCTGACGGGCGTCAACTGGTTTCGGCGGCGAGGGCTTCG	120	

CelF G	TCTTCAACCCCAACGGCATGGACATGCGCGGGTACTGGGGGATGATGGAGCAGATGAAGG	180	
CelF C	TCTTCAACCCCAACGGCATGGACATGCGCGGGTACTGGGGGATGATGGAGCAGATGAAGG	180	

CelF G	AGCTGGGCTTCAACACCATCCGCCTGCCGTGGAGCGACGCGGCTCTGGACGCCGACCGGG	240	
CelF C	AGCTGGGCTTCAACACCATCCGCCTGCCGTGGAGCGACGCGGCTCTGGACGCCGACCGGG	240	

CelF G	CGACGGGCATCGACACCTCCAAGAATCCCGACCTCGCGAATAAGTCCCCGCTTGAGGTCA	300	
CelF C	CGACGGGCATCGACACCTCCAAGAATCCCGACCTCGCGAATAAGTCCCCGCTTGAGGTCA	300	

(continua)

CelF G	TGGACAAGGTGATCGACTACGCCGGGCGCATCGGCATGAAGGTCATCCTCGACCACCATC	360
CelF C	TGGACAAGGTGATCGACTACGCCGGGCGCATCGGCATGAAGGTCATCCTCGACCACCATC	360

CelF G	GCTCCAGCGACGGTGCGTCGGCCAACGAGAACGGCCTGTGGTACGACGACAAGTACCCGG	420
CelF C	GCTCCAGCGACGGTGCGTCGGCCAACGAGAACGGCCTGTGGTACGACGACAAGTACCCGG	420

CelF G	AGTCGAAGATGATCGAAAACCTGGAAGATGCTGGCCGAGCGCTAC <u>A</u> AGGGCAACGACACGG	480
CelF C	AGTCGAAGATGATCGAAAACCTGGAAGATGCTGGCCGAGCGCTAC <u>T</u> AGGGCAACGACACGG	480
*****:*****		
CelF G	TGGTCGGCGCCGACCTGCACAACGAGCCGCACGGGCAGGCGACCTGGGGCGGTGGCGACA	540
CelF C	TGGTCGGCGCCGACCTGCACAACGAGCCGCACGGGCAGGCGACCTGGGGCGGTGGCGACA	540

CelF G	AGGCCACCGATTGGGCCTGGGCCGCGGAGCGCATCGGCAACGAGATCCAGTCGGTCAACA	600
CelF C	AGGCCACCGATTGGGCCTGGGCCGCGGAGCGCATCGGCAACGAGATCCAGTCGGTCAACA	600

CelF G	AGGATTGGCTGCTGCTGGTCGAGGGCGTCGAAATCCACCAGAATCAGTGGGAGTGGTGGG	660
CelF C	AGGATTGGCTGCTGCTGGTCGAGGGCGTCGAAATCCACCAGAATCAGTGGGAGTGGTGGG	660

CelF G	GCGGCAACCTGCGCGGCGCCAAGGACCGCCGATCGAATTCGATCATCCGGACAAGCTGG	720
CelF C	GCGGCAACCTGCGCGGCGCCAAGGACCGCCGATCGAATTCGATCATCCGGACAAGCTGG	720

CelF G	TCTACTCGGTCCACTCCTACGGCCCGTCCATCCACGAGATGGAGTGGTTCAAGGCCAGCG	780
CelF C	TCTACTCGGTCCACTCCTACGGCCCGTCCATCCACGAGATGGAGTGGTTCAAGGCCAGCG	780

CelF G	ACTACCCGAACAACCTGCCGGGCCAGTACACCGACAATTGGGGCTGGCTGCTGAAGGAAG	840
CelF C	ACTACCCGAACAACCTGCCGGGCCAGTACACCGACAATTGGGGCTGGCTGCTGAAGGAAG	840

CelF G	ACAAGGCGCCCGTGCTGGTCGGCGAGTTCGGCGGCAAGCTGGAGACCGATCAGGACAAGG	900
CelF C	ACAAGGCGCCCGTGCTGGTCGGCGAGTTCGGCGGCAAGCTGGAGACCGATCAGGACAAGG	900

CelF G	CCTACATGGCCAAGCTGATCGATTACATGAACGGCGACCTGGACGGCAACGGCACCAACG	960
CelF C	CCTACATGGCCAAGCTGATCGATTACATGAACGGCGACCTGGACGGCAACGGCACCAACG	960

CelF G	ATCTCGGCGCCGGCAAGGAAGGGGCGAGCTGGACCTACTGGTCGTGGAACCCGAACCTCGG	1020
CelF C	ATCTCGGCGCCGGCAAGGAAGGGGCGAGCTGGACCTACTGGTCGTGGAACCCGAACCTCGG	1020

CelF G	GCGACACCGGCGGCATCCTGAAGGACGATTGGCAGACGGTCGACCAGACCAAGTACAACG	1080
CelF C	GCGACACCGGCGGCATCCTGAAGGACGATTGGCAGACGGTCGACCAGACCAAGTACAACG	1080

CelF G	CCATCAAGGAAGGCCTGTTCGACGGCGGCAGCCCGCCAAGGCCGCCTGGGCTCCGGCCA	1140
CelF C	CCATCAAGGAAGGCCTGTTCGACGGCGGCAGCCCGCCAAGGCCGCCTGGGCTCCGGCCA	1140

(continua)

```

CelF G AGGCCCAGGCGCTGGGTGAGGGCGACATCAACGTCCTGTCCCACACGGTGACCCAGCCGC 1200
CelF C AGGCCCAGGCGCTGGGTGAGGGCGACATCAACGTCCTGTCCCACACGGTGACCCAGCCGC 1200
*****

CelF G AGGCGGCGGACCACAGCACCGCCGCGTCGGTCAACGCCGCGCTCGCCGCCAGACGGTGG 1260
CelF C AGGCGGCGGACCACAGCACCGCCGCGTCGGTCAACGCCGCGCTCGCCGCCAGACGGTGG 1260
*****

CelF G AGCTTCCGCCGGTCGATGCCGCGCACGATCCGTCGGCGCTGCACCACGACCCGCTGGACG 1320
CelF C AGCTTCCGCCGGTCGATGCCGCGCACGATCCGTCGGCGCTGCACCACGACCCGCTGGACG 1320
*****

CelF G TCCACACCGGCGCCTGATCGGGAGCCTGACAATCC 1355
CelF C TCCACACCGGCGCCTGATCGGGAGCCTTAGGATCC 1355
*****

```

Figura 17: Comparação de sequências do gene CelulaseF (CelF) presente no genoma de *A. brasilense* Sp245 com o inserto clonado presente em um dos clones obtidos. Programa ClustalW2 (LARKIN *et al.*, 2007). * para bases iguais, : para similares e . para pouco similares. As bases sublinhadas indicam a localização dos iniciadores utilizados na clonagem. Bases marcadas em cinza evidenciam o início e o fim do gene.

O primeiro clone analisado mostrou uma mutação inesperada na base 465 (mudança de uma adenina para timina), o que resultou na modificação K152STOP, ou seja, o códon para lisina 152 (códon AAG) foi trocado por um códon de parada (códon UAG, *amber nonsense mutation*), levando ao truncamento da proteína como mostrado na figura 18.

Alinhamento da sequência de aminoácidos do gene CelulaseF (CelF) presente no genoma (G) e no vetor clonado (C):

```

CelF G   MGIAKGFLHTDGNQIVDSSGQNVKLTGVNWFGGEGFVFNPNGMDMRGYWGMMEQMKELG  60
CelF C   MGIAKGFLHTDGNQIVDSSGQNVKLTGVNWFGGEGFVFNPNGMDMRGYWGMMEQMKELG  60
          *****

CelF G   NTIRLPWSDAALDADRATGIDTSKNPDLANKSPLEVMDKVIDYAGRIGMKVILDHHRSSD 120
CelF C   NTIRLPWSDAALDADRATGIDTSKNPDLANKSPLEVMDKVIDYAGRIGMKVILDHHRSSD 120
          *****

CelF G   GASANENGLWYDDKYPEskMIENWKMLAERYKGN DTVVGADLHNEPHGQATWGGGDKATD 180
CelF C   GASANENGLWYDDKYPEskMIENWKMLAERY •----- 151
          *****

CelF G   WAWAAERIGNEIQSVNKDWLLLVEGVEIHQNQWEWWGGNLRGAKDRPIEFDPDKLVYSV 240
CelF C   -----

CelF G   HSYGPSIHEMEWFKASDYPNNLPGQYTDNWGWLLKEDKAPVLVGEFGGKLETDQDKAYMA 300
CelF C   -----

CelF G   KLIDYMNGLDGNNGTNDLGAGKEGASWTYWSWNPNSGDTGGILKDDWQTVDTQTKYNAIKE 360
CelF C   -----

CelF G   GLFDGGSPAKAAWAPAKAQAALGEGDINVLSTVTQTPQAADHSTAASVNAALAAQTVELPP 420
CelF C   -----

CelF G   VDAAHDPsALHHDPLDVHTGA • 441
CelF C   -----

```

Figura 18: Comparação de sequências de aminoácidos do gene CelulaseF (CelF) presente no genoma de *A. brasilense* Sp245 com o inserto clonado (ClustalW2, LARKIN *et al.*, 2007). * para resíduos iguais. • para códon de parada.

A análise da sequência presente no segundo clone (figura 19) apontou 2 mudanças de bases, além daquelas inseridas nos iniciadores. As modificações sofridas levam a substituições K177E e D268G.

Alinhamento do gene CelulaseF (CelF) presente no genoma (G) e o clonado (C):

```

CelF G   GGCTGAGCACAAATGGGCATCGCCAAGGGCTTCCTGCACACGGACGGCAACCAGATCGTCG  60
CelF C   GGCTGAGACATATGGGCATCGCCAAGGGCTTCCTGCACACGGACGGCAACCAGATCGTCG  60
          *****...:*****

CelF G   ATTCGTCGGGTCAGAACGTCAAGCTGACGGGCGTCAACTGGTTCGGCGGCGAGGGCTTCG 120
CelF C   ATTCGTCGGGTCAGAACGTCAAGCTGACGGGCGTCAACTGGTTCGGCGGCGAGGGCTTCG 120
          *****

```

(continua)

CelF	G	TCTTCAACCCCAACGGCATGGACATGCGCGGGTACTGGGGGATGATGGAGCAGATGAAGG	180
CelF	C	TCTTCAACCCCAACGGCATGGACATGCGCGGGTACTGGGGGATGATGGAGCAGATGAAGG	180

CelF	G	AGCTGGGCTTCAACACCATCCGCCTGCCGTGGAGCGACGCGGCTCTGGACGCCGACCGGG	240
CelF	C	AGCTGGGCTTCAACACCATCCGCCTGCCGTGGAGCGACGCGGCTCTGGACGCCGACCGGG	240

CelF	G	CGACGGGCATCGACACCTCCAAGAATCCCGACCTCGCGAATAAGTCCCCGCTTGAGGTCA	300
CelF	C	CGACGGGCATCGACACCTCCAAGAATCCCGACCTCGCGAATAAGTCCCCGCTTGAGGTCA	300

CelF	G	TGGACAAGGTGATCGACTACGCCGGGCGCATCGGCATGAAGGTCATCCTCGACCACCATC	360
CelF	C	TGGACAAGGTGATCGACTACGCCGGGCGCATCGGCATGAAGGTCATCCTCGACCACCATC	360

CelF	G	GCTCCAGCGACGGTGCGTCGGCCAACGAGAACGGCCTGTGGTACGACGACAAGTACCCGG	420
CelF	C	GCTCCAGCGACGGTGCGTCGGCCAACGAGAACGGCCTGTGGTACGACGACAAGTACCCGG	420

CelF	G	AGTCGAAGATGATCGAAAACCTGGAAGATGCTGGCCGAGCGCTACAAGGGCAACGACACGG	480
CelF	C	AGTCGAAGATGATCGAAAACCTGGAAGATGCTGGCCGAGCGCTACAAGGGCAACGACACGG	480

CelF	G	TGGTCGGCGCCGACCTGCACAACGAGCCGCACGGGCAGGCGACCTGGGGCGGTGGCGAC <u>A</u>	540
CelF	C	TGGTCGGCGCCGACCTGCACAACGAGCCGCACGGGCAGGCGACCTGGGGCGGTGGCGAC <u>G</u>	540

CelF	G	AGGCCACCGATTGGGCCTGGGCCGCGGAGCGCATCGGCAACGAGATCCAGTCGGTCAACA	600
CelF	C	AGGCCACCGATTGGGCCTGGGCCGCGGAGCGCATCGGCAACGAGATCCAGTCGGTCAACA	600

CelF	G	AGGATTGGCTGCTGCTGGTCGAGGGCGTCGAAATCCACCAGAATCAGTGGGAGTGGTGGG	660
CelF	C	AGGATTGGCTGCTGCTGGTCGAGGGCGTCGAAATCCACCAGAATCAGTGGGAGTGGTGGG	660

CelF	G	GCGGCAACCTGCGCGGGCGCCAAGGACCGCCCGATCGAATTCGATCATCCGGACAAGCTGG	720
CelF	C	GCGGCAACCTGCGCGGGCGCCAAGGACCGCCCGATCGAATTCGATCATCCGGACAAGCTGG	720

CelF	G	TCTACTCGGTCCACTCCTACGGCCCGTCCATCCACGAGATGGAGTGGTTCAAGGCCAGCG	780
CelF	C	TCTACTCGGTCCACTCCTACGGCCCGTCCATCCACGAGATGGAGTGGTTCAAGGCCAGCG	780

CelF	G	ACTACCCGAACAACCTGCCGGGCCAGTACACCG <u>A</u> CAATTGGGGCTGGCTGCTGAAGGAAG	840
CelF	C	ACTACCCGAACAACCTGCCGGGCCAGTACACCG <u>G</u> CAATTGGGGCTGGCTGCTGAAGGAAG	840

CelF	G	ACAAGGCGCCCGTGCTGGTCGGCGAGTTCGGCGGCAAGCTGGAGACCGATCAGGACAAGG	900
CelF	C	ACAAGGCGCCCGTGCTGGTCGGCGAGTTCGGCGGCAAGCTGGAGACCGATCAGGACAAGG	900

CelF	G	CCTACATGGCCAAGCTGATCGATTACATGAACGGCGACCTGGACGGCAACGGCACCAACG	960
CelF	C	CCTACATGGCCAAGCTGATCGATTACATGAACGGCGACCTGGACGGCAACGGCACCAACG	960

(continua)

CelF	G	ATCTCGGCGCCGGCAAGGAAGGGGCGAGCTGGACCTACTGGTCGTGGAACCCGAACCTCGG	1020
CelF	C	ATCTCGGCGCCGGCAAGGAAGGGGCGAGCTGGACCTACTGGTCGTGGAACCCGAACCTCGG	1020

CelF	G	GCGACACCGGCGGCATCCTGAAGGACGATTGGCAGACGGTCGACCAGACCAAGTACAACG	1080
CelF	C	GCGACACCGGCGGCATCCTGAAGGACGATTGGCAGACGGTCGACCAGACCAAGTACAACG	1080

CelF	G	CCATCAAGGAAGGCCTGTTTCGACGGCGGCAGCCCGGCCAAGGCCGCTGGGCTCCGGCCA	1140
CelF	C	CCATCAAGGAAGGCCTGTTTCGACGGCGGCAGCCCGGCCAAGGCCGCTGGGCTCCGGCCA	1140

CelF	G	AGGCCCAGGCGCTGGGTGAGGGCGACATCAACGTCCTGTCCCACACGGTGACCCAGCCGC	1200
CelF	C	AGGCCCAGGCGCTGGGTGAGGGCGACATCAACGTCCTGTCCCACACGGTGACCCAGCCGC	1200

CelF	G	AGGCGGCGGACCACAGCACC GCCGCTCGGTCAACGCCGCGCTCGCCGCCAGACGGTGG	1260
CelF	C	AGGCGGCGGACCACAGCACC GCCGCTCGGTCAACGCCGCGCTCGCCGCCAGACGGTGG	1260

CelF	G	AGCTTCCGCCGGTCGATGCCGCGCACGATCCGTCGGCGCTGCACCACGACCCGCTGGACG	1320
CelF	C	AGCTTCCGCCGGTCGATGCCGCGCACGATCCGTCGGCGCTGCACCACGACCCGCTGGACG	1320

CelF	G	TCCACACCGGCGCCTGATCGGGAGCCTGACAATCC	1355
CelF	C	TCCACACCGGCGCC <u>TGAT</u> CGGGAGCCTTAGGATCC	1355
		***** * . ****	

Figura 19: Comparação de seqüências do gene CelulaseF (CelF) presente no genoma de *A. brasilense* Sp245 com o segundo inserto clonado (ClustalW2, LARKIN *et al.*, 2007). * para bases iguais, : para similares e . para pouco similares. As bases sublinhadas indicam a localização dos iniciadores utilizados na clonagem. Bases marcadas em cinza evidenciam o início e o fim do gene.

De maneira similar aos genes anteriores, o inserto presente no segundo clone analisado (clonado no vetor pTZ57R/T, denominado pMS_zCelF) foi ligado ao vetor pET28a, mesmo apresentando as mutações. Ao final da transformação em *E. coli* DH10B, os clones que apresentaram o inserto foram conservados em glicerol 50% a -20°C, e um plasmídeo foi selecionado aleatoriamente e passou a ser chamado de pMS_CelF.

5.2.4 Clonagem do gene CelulaseG de *A. brasilense* Sp245

O gene CelulaseG (linha 9 na tabela 9, p. 42) codifica pra uma proteína pertencente à classe GH5 das glicosil-hidrolases. Os clones que possuem esse inserto em seus plasmídeos devem gerar bandas em torno de 2265 pb após a PCR de colônia.

O fragmento de PCR contendo o gene com o uso do DNA genômico de *A. brasilense* Sp245 foi ligado ao vetor pTZ57R/T e o produto de ligação foi

transformado em *E. coli* DH10B, gerando 8 colônias brancas. Esses clones foram submetidos ao teste de PCR de colônia e os resultados estão mostrados a seguir (figura 20).

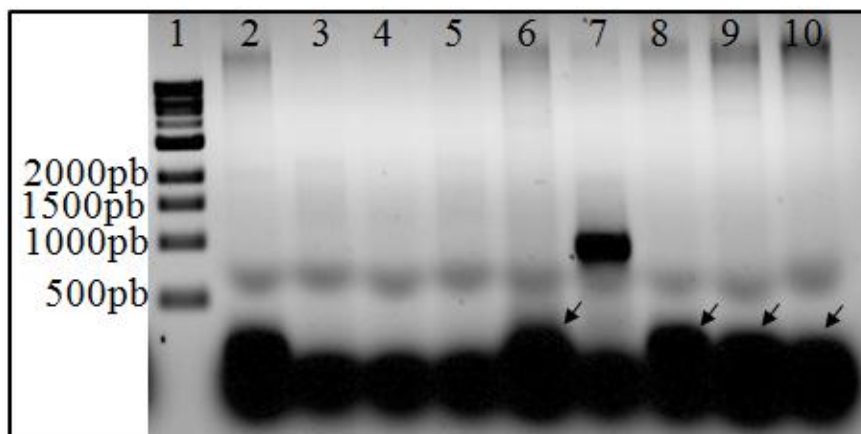


Figura 20: Identificação de colônias transformantes contendo o vetor pTZ57R/T + gene CelulaseG. Eletroforese em gel de agarose 1% mostrando o resultado da PCR de colônia utilizando iniciadores M13 universal e reverso. Poço 1: marcador de tamanho molecular. Poço 2: controle negativo. Poços 3 a 10: colônias brancas obtidas após transformação. As setas indicam bandas que podem ser similares à presente no controle negativo.

Das 8 colônias analisadas, 4 apresentaram uma possível banda abaixo de 500 pb, se desconsiderarmos as bandas circulares presentes em todas as amostras (provavelmente por excesso de iniciadores adicionados ao ensaio), similar àquela presente no controle negativo, e um clone mostrou uma banda inespecífica em torno de 1000 pb. Para confirmar a ausência ou presença de inserto nos plasmídeos das 3 colônias brancas restantes (poços 3, 4 e 5), seus materiais genéticos foram purificados e analisados por digestão com as enzimas *EcoRI* e *HindIII*, seguida de eletroforese em gel de agarose. Os produtos de restrição esperados eram 3000 pb (vetor aberto) e 2200 pb (inserto). O resultado está mostrado na figura 21

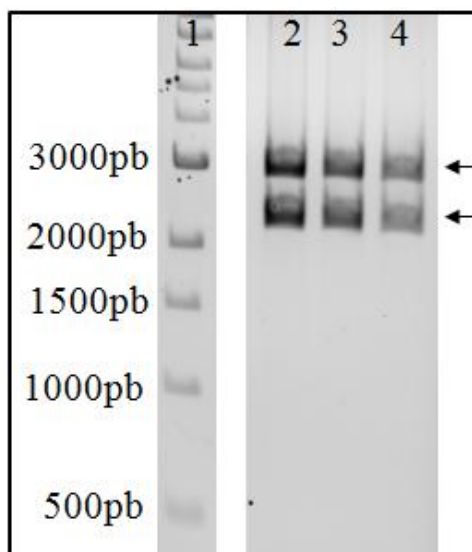


Figura 21: Análise de restrição com as enzimas *EcoRI* e *HindIII* de 3 plasmídeos derivados do vetor pTZ57R/T+CelulaseG, em gel de agarose 1%. Poço 1: marcador de tamanho molecular. Poços 2 a 4: plasmídeos de 3 clones transformantes após ligação do vetor pTZ57R/T com gene CelulaseG. As setas mostram as duas bandas esperadas após a restrição: uma com ~3000 pb (vetor aberto) e outra com ~2200 pb (gene).

Mesmo não apresentando a banda esperada no teste de PCR de colônia, os 3 clones analisados aparentam ter o fragmento do gene CelulaseG em seus plasmídeos. Aparentemente, a dificuldade de amplificação vista na etapa inicial de clonagem se manteve na técnica de PCR de colônia, sugerindo então que o gene em questão pode apresentar estruturas secundárias tanto no genoma quanto no plasmídeo após clonagem, dificultando a atividade da DNA polimerase.

Para confirmar e identificar possíveis mutações, os plasmídeos foram sequenciados utilizando iniciadores M13 universal e reverso, porém, devido ao tamanho do gene CelulaseG, não foi obtida a sua sequência completa. Novos experimentos deverão ser conduzidos para possibilitar a sua obtenção. Mesmo assim, foram analisadas as sequências parciais obtidas das extremidades do gene (em torno de 1000 pb, aproximadamente metade do gene) e não houve a identificação de mutações nestas regiões.

Em paralelo, o inserto de um dos 3 clones foi purificado e recebeu o nome de pMS_zCelG. Este teve seu inserto purificado e, posteriormente, foi ligado ao vetor de expressão pET28a. Depois da transformação em *E. coli* DH10B, o inserto analisado gerou colônias transformantes. O plasmídeo obtido foi chamado de pMS_CelG.

5.3 EXPRESSÃO E PURIFICAÇÃO DOS GENES CLONADOS

Embora os clones dos 4 genes codificando prováveis glicosil-hidrolases apresentassem mutações, as proteínas foram superexpressas, purificadas e utilizadas em ensaios de atividade.

Para analisar a solubilidade das enzimas realizou-se a expressão em *E. coli* BL21 (DE3) em pequena escala (20 mL de cultivo).

No primeiro teste de expressão, utilizaram-se apenas 3 horas de indução a 37°C. As células foram coletadas por centrifugação e suspensas em tampão de sonicação (50 mM Tris-HCl pH 7,5; 50 mM de NaCl) (figura 22).

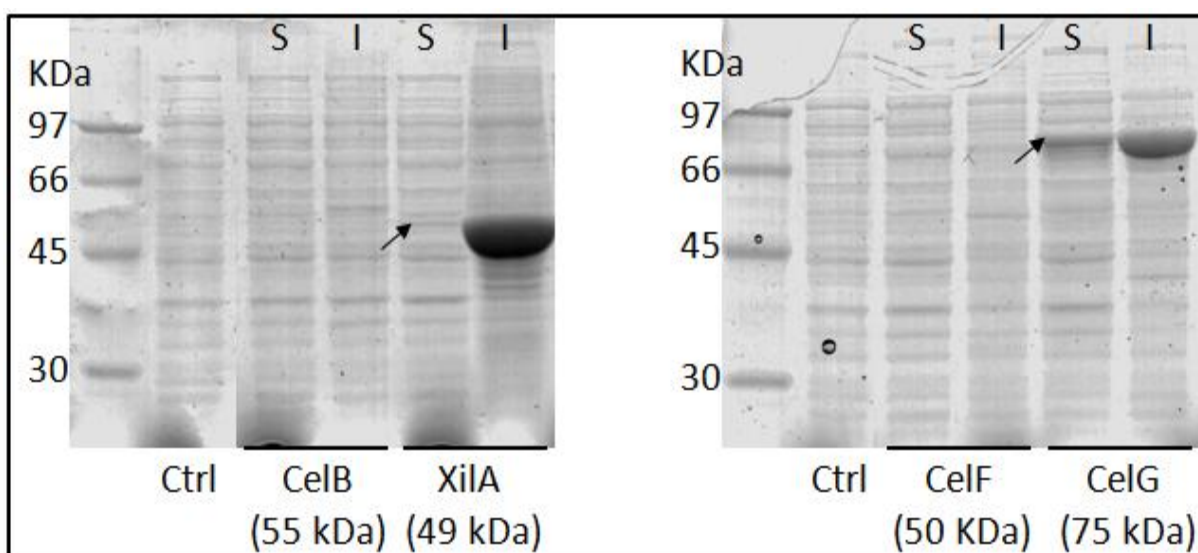


Figura 22: Perfil eletroforético em gel de poli-acrilamida (10%) de extrato de *E. coli* BL21 (DE3) expressando as proteínas indicadas (CelB para a CelulaseB, XilA para XilanaseA, CelF para CelulaseF e CelG para CelulaseG). As culturas foram induzidas pela adição de IPTG 0,3 mM e incubadas a 37°C por 3 horas. As setas mostram possíveis bandas relativas à proteína expressa. A cultura controle ("Ctrl") é extrato celular de *E. coli* BL21 (DE3), cultivada nas mesmas condições do experimento, contendo o vetor pET28a vazio. S indica fração solúvel, e I, fração insolúvel. Os números entre parênteses (na parte inferior da figura) indicam os tamanhos das proteínas.

Os primeiros resultados mostraram que a proteína XilanaseA é bem expressa, mas encontra-se na fração insolúvel. Resultado similar foi obtido para a proteína CelulaseG, que também apresentou expressão elevada, embora tenha apresentado uma solubilidade maior que a XilanaseA. Já para as proteínas CelulaseB e CelulaseF não foram superexpressas nestas condições.

Um segundo experimento de indução foi realizado modificando as condições de indução (18°C por 16 horas) e o tampão de sonicação (50 mM de Tris-HCl e 50 mM de NaCl, pH 7,5). Após adição de IPTG 0,3 mM as culturas foram incubadas por 16 horas, a 18°C, as células foram então coletadas e lisadas por sonicação. Os extratos foram analisados por eletroforese em poli-acrilamida (figura 23).

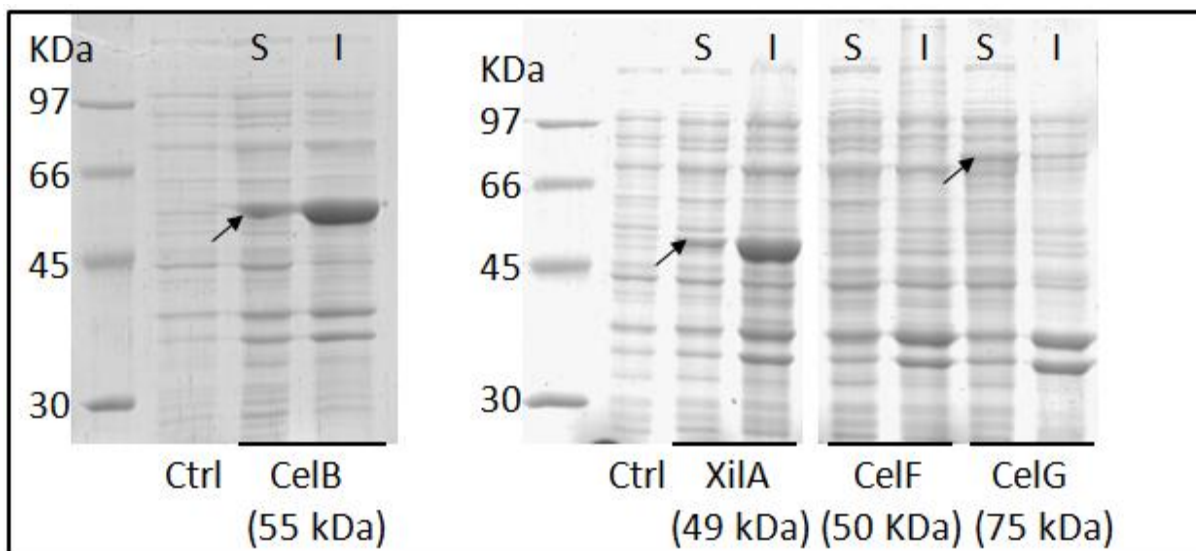


Figura 23: gel de poli-acrilamida (10%) mostrando o resultado do segundo teste de indução feito a 18°C, *overnight*. As setas apontam para bandas diferenciais em relação ao controle (extrato celular de *E. coli* BL21 (DE3), cultivada nas mesmas condições contendo plasmídeo pET28a vazio). Legenda: S significa fração solúvel, e I, fração insolúvel. Os números entre parênteses (na parte inferior da figura) indicam os tamanhos das proteínas (CelB para a CelulaseB, XilA para XilanaseA, CelF para CelulaseF e CelG para CelulaseG).

Em baixas temperaturas, a proteína CelulaseB apresentou uma expressão elevada, onde a maior concentração proteica está presente na fração insolúvel. A proteína XilanaseA foi bem expressa nesta condição novamente, embora não tão bem quanto a 37°C, mas uma maior concentração da proteína se encontrava na fase solúvel do extrato. Um resultado similar foi obtido para a proteína CelulaseG. Não se detectou expressão da proteína CelulaseF nesta condição e por isto esta proteína não foi utilizada nos experimentos subsequentes.

Como a indução a 18°C gerou um quantidade maior de proteína solúvel em todos os casos, esta condição foi utilizada para purificar as proteínas XilanaseA e CelulaseG primeiramente. Após as etapas de indução, os extratos foram obtidos e as proteínas purificadas por cromatografia de afinidade em coluna *HiTrap™ Chelating* como descrito em Material e Métodos.

A figura 24 mostra as eletroforeses de amostras de cada etapa de purificação de XilanaseA. A proteína foi eluída nas concentrações de 300 e 500 mM de imidazol, sendo que a fração eluída com 500 mM de imidazol aparenta estar mais pura. Assim, as frações eluídas com 300 mM e 500 mM de imidazol foram unidas e dialisadas, gerando as alíquotas XilA-I (0,28 µg de proteína /µL) e XilA-II (0,11 µg de proteína/µL), respectivamente. Notou-se formação de precipitado após a diálise.

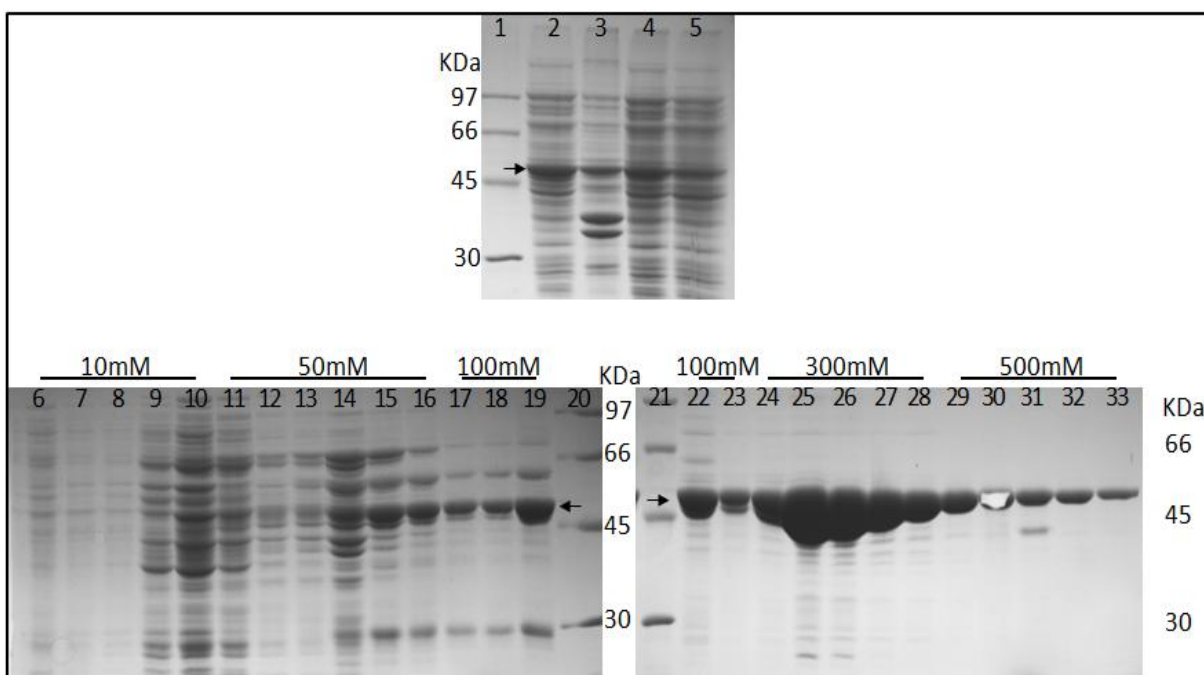


Figura 24: géis de poli-acrilamida (10%) contendo as amostras obtidas na expressão e purificação da proteína XilanaseA. Poços 1, 20 e 21: marcadores de peso molecular. 2: extrato bruto. 3: fração insolúvel (*pellet*). 4: fração não ligada. 5: lavagem da coluna. Alíquotas das frações eluídas com 10 mM (poços 6-10), 50 mM (11-16), 100 mM (17-19 e 22-23), 300 mM (24-28) e 500 mM (29-33) de imidazol. A seta indica a localização da proteína XilanaseA nas alíquotas, e a banda da proteína na linha 30 foi retirada para confirmação em MALDI-TOF-TOF.

Para garantir que a proteína purificada era de fato a XilanaseA, coletou-se a banda da proteína na fração eluída com 500 mM de imidazol. Esta foi recortada do gel e submetida ao protocolo de identificação de proteína por espectrometria de massas (maiores detalhes na seção 4.6.4). O resultado (tabela 10) confirmou se tratar da proteína XilanaseA.

Tabela 10: Identificação da proteína XilanaseA através de espectrometria de massa MS e MS/MS.

Método	Cobertura de sequência	Escore proteína	Número de acesso	Peptídeo observado/ identificados ¹	Seq. peptídeo (ion score) ²
MS	44%	158	AZOBR_p1110166*	26/17	-
MS/MS	3%	39	AZOBR_p1110166*	-	GPEWLDPFFHHIR (39)

1- número de peptídeos buscados e pareados nos dados de impressão digital de peptídeos (PMF); 2- íon selecionado para fragmentação no método MS/MS; *gene da XilanaseA no genoma de *Azospirillum brasilense* Sp245.

A proteína CelulaseG foi purificada e os perfis eletroforéticos de amostras das etapas de purificação estão na figura 25.

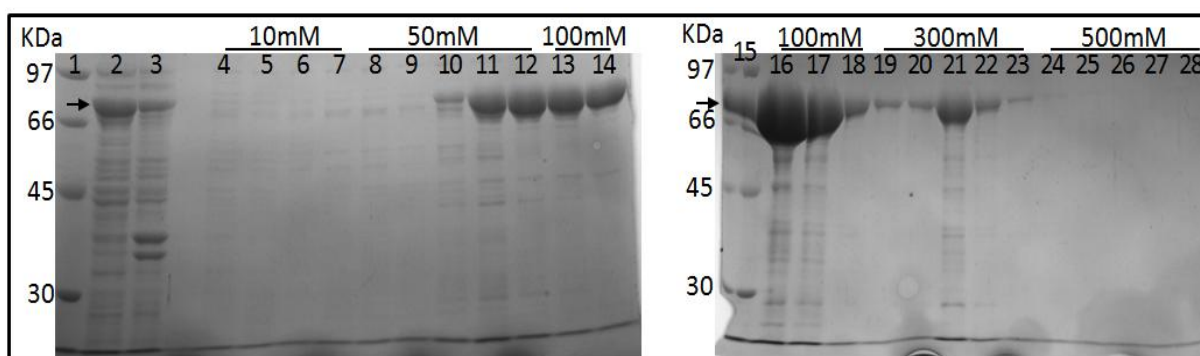


Figura 25: géis de poli-acrilamida (10%) contendo as amostras obtidas na expressão e purificação da proteína CelulaseG. 1 e 15: marcadores de peso molecular. 2: extrato bruto. 3: fração insolúvel (*pellet*). Alíquotas das frações eluídas com 10 mM (4-7), 50 mM (8-12), 100 mM (13-14 e 16-18), 300 mM (19-23) e 500 mM (24-28) de imidazol. A seta indica a proteína CelulaseG nas frações. O marcador de peso molecular na linha 15 foi contaminado pela alíquota presente na linha 16, por isso há uma banda entre os marcadores de 97 e 66 kDa.

O perfil de eluição para a proteína CelulaseG se iniciou com 50 mM de imidazol e foi até 300 mM, sendo que a maior quantidade da proteína foi eluída com 100 mM de imidazol. Aparentemente parte das amostras nas linhas 18, 19 e 20 foram perdidas durante a aplicação das amostras no gel. Assim, as alíquotas foram unidas e dialisadas, formando as frações CelG-1 (linha 11 ao 14; 0,16 µg de proteína/µL), CelG-2 (16 ao 18; 1,31 µg de proteína /µL) e CelG-3 (19 ao 22; 0,21 µg de proteína /µL).

A identidade de CelulaseG também foi confirmada por espectrometria de massas MS/MS (tabela 11).

Tabela 11: Identificação da proteína CelulaseG através de espectrometria de massa MS/MS.

Método	Cobertura de sequência	Escore proteína	Número de acesso	Seq. peptídeo (ion score) ¹
MS/MS	1%	60	AZOBR_p470008*	DIWTDAGWIVQK (60)

1- íon selecionado para fragmentação no método MS/MS; *gene da CelulaseG no genoma de *Azospirillum brasilense* Sp245.

Como as proteínas purificadas não apresentaram atividade (será discutido posteriormente), optou-se por realizar, em primeiro lugar, testes de atividade no extrato bruto contendo a proteína CelulaseB.

5.4 ENSAIOS DE ATIVIDADE ENZIMÁTICA

Para determinar os prováveis substratos das proteínas expressas foram realizadas comparações com o banco de dados CAZY (*Carbohydrate-active enzymes*), que é um banco de dados de glicosil-hidrolases. Foram analisados quais substratos podem ser degradados pelas classes das enzimas estudadas e, através da ferramenta BlastX (ALTSCHUL *et al.*, 1997), buscou-se encontrar similaridade com enzimas que degradem tais substratos. Os resultados obtidos estão presentes na tabela 12.

Tabela 12: Identificação de possíveis substratos para as enzimas obtidas. Compilação dos dados presentes no banco de dados CAZY e os resultados de busca do BlastX.

Gene	Classe	Possíveis atividades (% de identidade)	Substrato
CelulaseB	GH1	β -glucosidase (52%), β -galactosidase (54%)	celobiose, lactose
CelulaseG	GH5	quitosanase (NE), endo-glucanase (54%), exo-glucanase (NE), quitinase (52%)	quitosana, celulose, quitina
XilanaseA	GH10	β -xilanase (28%)	xilana

NE- não encontrado enzimas com a atividade descrita após comparação usando BlastX.

Os ensaios de atividade foram feitos inicialmente com a proteína XilanaseA, para a qual se utilizou como substrato 1% de xilana em tampão (100 mM de Tris-HCl pH7,5 e 100 mM de KCl) e 30°C. Após várias tentativas, não se detectou evidência de hidrólise utilizando o método de açúcar redutor, mesmo aumentando a quantidade de enzima adicionada (de 5 a 70 μ g), o tempo de reação (1 e 16 horas) e modificando o pH (5; 7; 8,8 e 10). Tentou-se também adicionar à reação 100 mM de Ca^{2+} ou de Mg^{2+} , como possíveis co-fatores, mas ainda não se detectou liberação de açúcar redutor pelo teste do DNS.

Já para a enzima CelulaseG, utilizou-se como substrato 1% de CMC (carboximetilcelulose sódica), 1% de quitosana ou 1% de quitina, todos em tampão 100 mM de Tris-HCl pH7,5 e 100 mM de KCl. Não se detectou nenhuma atividade mesmo utilizando 250 μ g de enzima, 16 horas de incubação com substrato ou variando o pH do tampão (5; 7,5; 8,8 ou 10).

Uma hipótese levantada foi que as enzimas XilanaseA e CelulaseG poderiam ter perdido suas atividades durante as etapas de purificação. Assim, as 3 enzimas (CelulaseB, CelulaseG e XilanaseA) foram expressas e os cultivos foram sonificados em tampão de sonicação. Os extratos brutos foram quantificados pelo método de

Bradford e juntou-se 300 µg de massa protéica de cada alíquota a 500 µL de tampão 100mM Tris-HCl pH 7,5 contendo substrato (1% de CMC ou 1% de xilana). Mesmo assim, o teste do DNS não identificou nenhum açúcar redutor formado, indicando que as enzimas estão inativas nestas condições ou que o substrato fornecido não é atacado por elas.

Outra possibilidade é que as mutações resultantes da PCR podem ter afetado negativamente a atividade da enzima. Para verificar esta possibilidade, foram realizadas análises da estrutura secundária para o gene com e sem mutação e comparações da sequência de aminoácidos das proteínas de *A. brasilense* Sp245 e proteínas homólogas. Para a proteína CelulaseB, as mutações L138P e N437D modificaram resíduos que podem ser importantes para a enzima, uma vez que a comparação feita (figura 26) revelou que estes resíduos de leucina (L) e asparagina (N) são conservados em enzimas de outros microrganismos.

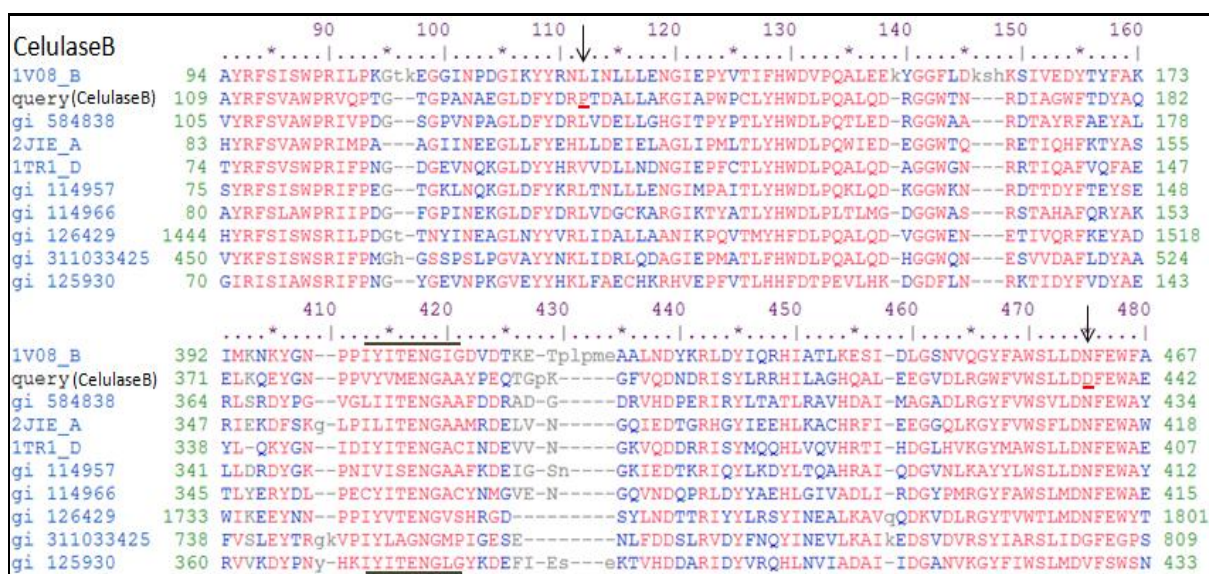


Figura 26: alinhamento entre sequências similares à proteína CelulaseB utilizando o programa CD-Search (MARCHLER-BAUER *et al.*, 2011). As flechas e os sublinhados vermelhos indicam as mutações ocorridas no processo de clonagem. As barras em cinza indicam a possível localização da região catalítica para a família GH1 das glicosil-hidrolases (com base nos dados gerados pelo programa PROSITE - ExPASy). Comparação feita entre β-glucosidades/β-galactosidades de: milho (1V08_b; organismo eucarioto), *Thermobispora bispora* (584838; organismo procarioto), *Paenibacillus polymyxa* (2JIE_A e 1TR1_D; organismo procarioto), *Clostridium thermocellum* (114957; organismo procarioto), *Agrobacterium sp.* (114966; organismo procarioto), *Oryctolagus cuniculus* (136429; organismo eucarioto), *Homo sapiens* (311033425; organismos eucarioto), *Staphylococcus aureus* (125930; organismo procarioto).

Segundo o programa PsiPred (JONES, 1999), as mutações provavelmente não alteraram a estrutura secundária da enzima na região onde elas ocorreram (figura 27).

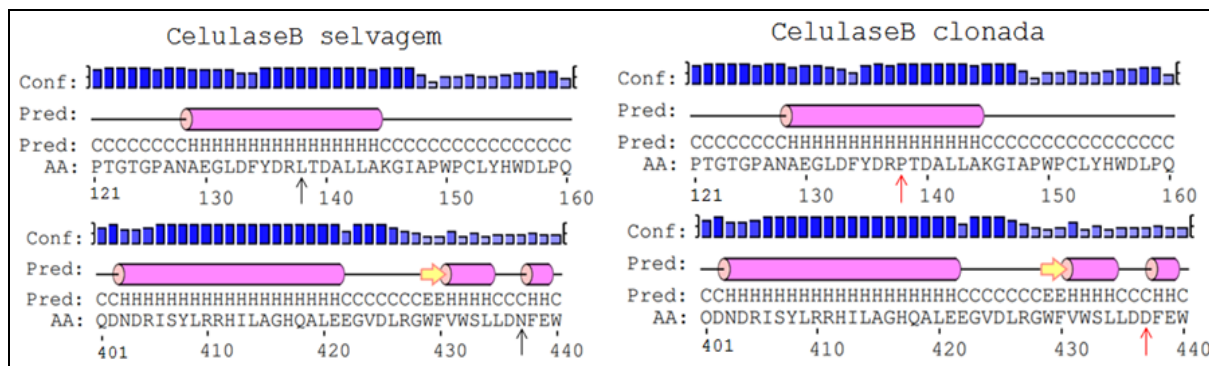


Figura 27: Predição da estrutura secundária da proteína CelulaseB (selvagem e clonada) utilizando o programa PsiPred (JONES, 1999). Estão mostradas apenas as regiões onde ocorreram as mutações. As flechas indicam os resíduos que sofreram mutação. Conf: grau de confiabilidade calculado pelo programa para a estrutura secundária predita. Pred: predição de estrutura, onde C representa resíduos que contribuem para a formação de estrutura tipo alça, E, resíduos que contribuem para a formação de folha β e H, resíduos que contribuem para a formação de α -hélice. AA: resíduos de aminoácidos analisados. O cilindro rosa corresponde a uma possível estrutura α -hélice. As setas amarelas correspondem a folha β .

Por outro lado as mutações em XilanaseA (A41V e L56P) encontram-se em uma região de baixa conservação entre as proteínas da mesma classe de glicosil-hidrolases (figura 28). A posição 41 pode conter A, F, T, L ou C, sugerindo que esta modificação não é importante. Já a posição 56 é mais conservada, sendo preferencialmente ocupada por resíduo de aminoácido com cadeia lateral grande e apolar (L, V ou I). A substituição por P pode levar a disrupção da estrutura secundária levando a perda da atividade. A análise feita pelo programa PsiPred corrobora com a hipótese levantada anteriormente, onde ocorre a desestruturação de uma possível α -hélice nos dois pontos mutados (figura 29).



Figura 28: alinhamento da sequência de XilanaseA obtida e de proteínas similares utilizando o programa CD-Search (MARCHLER-BAUER *et al.*, 2011). As setas e os sublinhados vermelhos indicam as mutações ocorridas no processo de clonagem. As barras em cinza indicam a possível localização da região catalítica para a família GH10 das glicosil-hidrolases (com base nos dados gerados pelo programa PROSITE - ExpASY). Comparação feita entre xilanases de: *Thermotoga marítima* (2494333 e 15642845; organismo procarioto), *Agrobacterium tumefaciens* (15889644; organismo procarioto), *Bacillus halodurans* (139867; organismo procarioto), *Clostridium acetobutylicum* (15004757 e 15004819; organismo procarioto), *Caulobacter crescentus* (16127035 e 16127272; organismo procarioto) e *Nostoc sp.* (17231940; organismo procarioto).

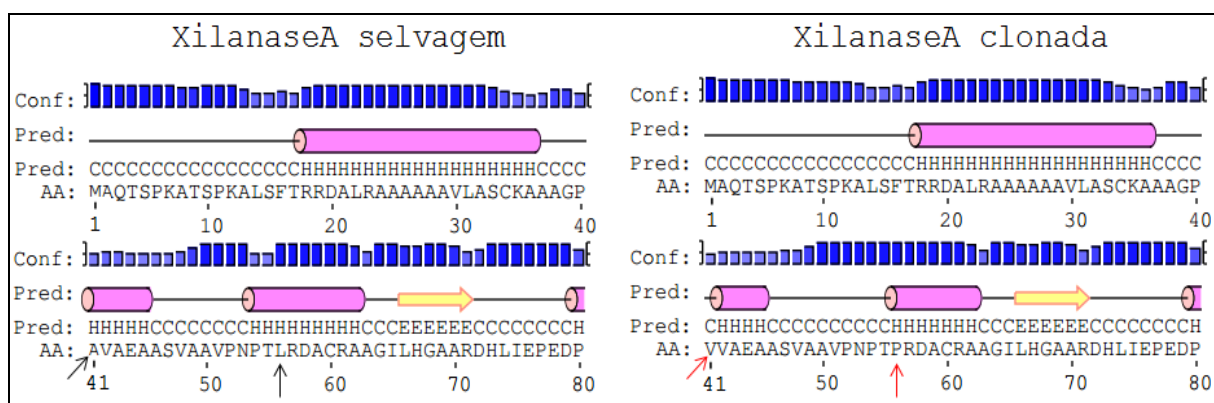


Figura 29: Predição da estrutura secundária de XilanaseA (selvagem e clonada) utilizando o programa PsiPred (JONES, 1999). Estão mostradas apenas a região N-terminal, onde ocorreram as mutações. As flechas indicam os resíduos que sofreram mutação. Conf: grau de confiabilidade calculado pelo programa para a estrutura secundária predita. Pred: predição de estrutura, onde C representa resíduos que contribuem para a formação de estrutura tipo alça, E, resíduos que contribuem para a formação de folha β e H, resíduos que contribuem para a formação de α -hélice. AA: resíduos de aminoácidos analisados. O cilindro rosa corresponde a uma possível estrutura α -hélice. As setas amarelas correspondem a folha β .

Foi analisado também se as mutações presentes em CelulaseF poderiam influenciar a sua atividade. As modificações identificadas estão em regiões pouco conservadas (figura 30), mas nota-se que os aminoácidos presentes nas outras

sequências são polares ou carregados, ao contrário do aminoácido alterado (glicina – apolar). Segundo a análise do PsiPred, a mutação D268G acarretou na perda de uma região α -hélice (figura 31), o que pode explicar a inatividade da enzima. Mesmo assim, a cauda de histidina fusionada, ausência de algum co-fator ou o uso de substrato incorreto também podem estar relacionados com a inatividade hidrolítica nos ensaios realizados.

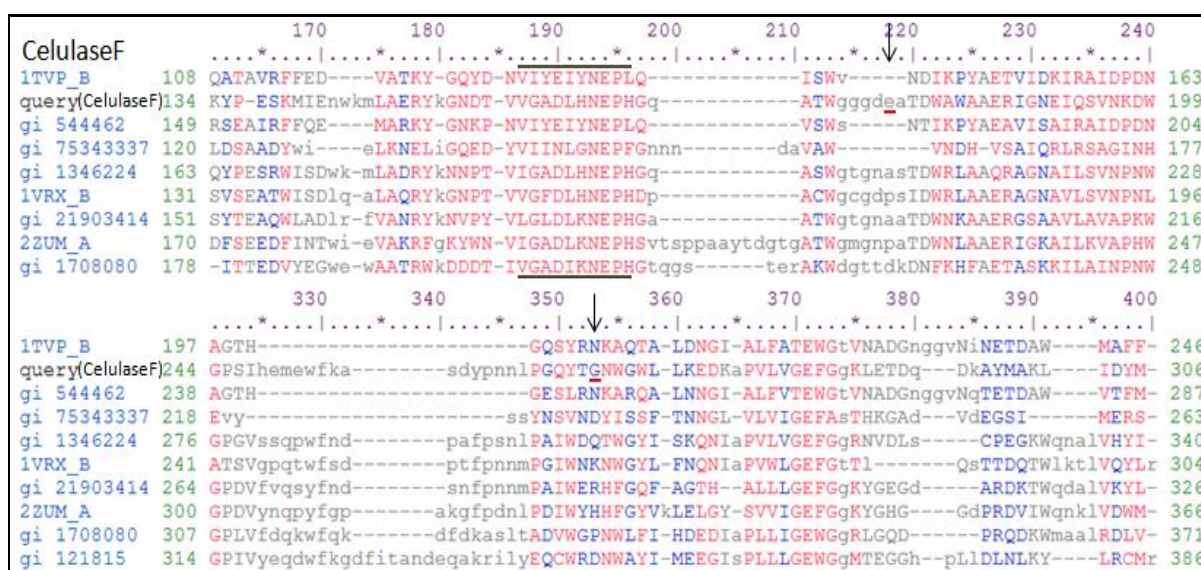


Figura 30: alinhamento entre sequências similares à enzima CelulaseF utilizando o programa CD-Search (MARCHLER-BAUER *et al.*, 2011). As flechas e os sublinhados vermelhos indicam as mutações ocorridas no processo de clonagem. As barras em cinza indicam a possível localização da região catalítica para a família GH5 das glicosil-hidrolases (com base nos dados gerados pelo programa PROSITE - ExPASy). Comparação feita entre: *Pseudoalteromonas haloplanktis* (1TVP_B, organismo procarioto), *Dickeya dadantii* 3937 (544462, organismo procarioto), *Vibrio* sp. (75343337, organismo procarioto), *Paenibacillus polymyxa* (1346224, organismo procarioto), *Acidothermus cellulolyticus* (1VRX_B, organismo procarioto), *Xanthomonas campestris* (21903414, organismo procarioto), *Pyrococcus horikoshii* OT3 (2ZUM_A, organismo procarioto), *Cellulomonas fimi* (1708080, organismo procarioto), *Clostridium thermocellum* (121815, organismo procarioto).

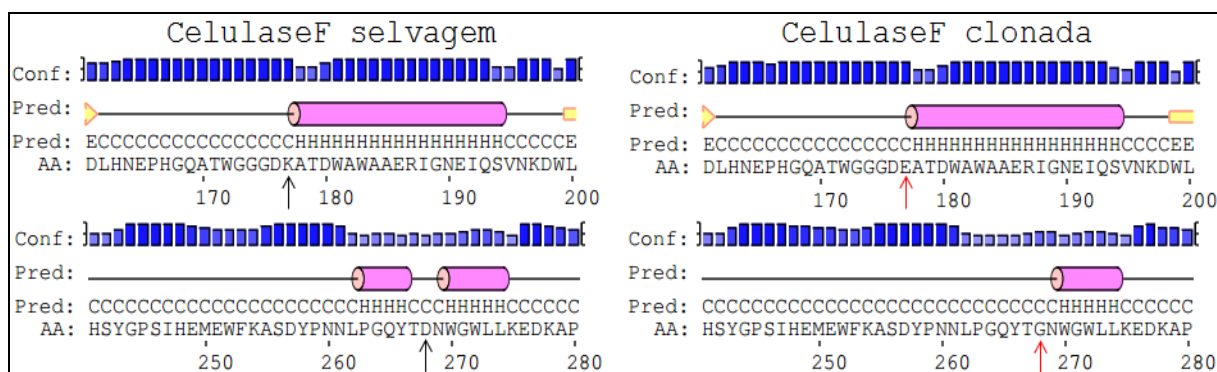


Figura 31: Predição da estrutura secundária para CelulaseF (selvagem e clonada) utilizando o programa PsiPred (JONES, 1999). Estão mostradas apenas as regiões onde ocorreram as mutações. As flechas indicam os resíduos que sofreram mutação. Conf: grau de confiabilidade calculado pelo programa para a estrutura secundária predita. Pred: predição de estrutura, onde C representa resíduos que contribuem para a formação de estrutura tipo alça, E, resíduos que contribuem para a formação de folha β e H, resíduos que contribuem para a formação de α -hélice. AA: resíduos de aminoácidos analisados. O cilindro rosa corresponde a uma possível estrutura α -hélice. As setas amarelas correspondem a folha β .

Como o gene da proteína CelulaseG ainda não foi sequenciado completamente, é possível que possa ter acontecido alguma mutação na região interna de sua sequência ou a falta de algum co-fator que impediu a sua atividade contra os substratos testados (quitosana, quitina e celulose).

5.5 CULTIVO DE *A. BRASILENSE* EM MEIO CONTENDO APENAS CMC E XILANA COMO FONTE DE CARBONO

Para confirmar se *Azospirillum brasilense* realmente apresenta enzimas capazes de degradar CMC (carboximetilcelulose) ou xilana foi utilizada a estirpe SP245 de *A. brasilense* e também a estirpe FP2. Para isto as estirpes foram inoculadas em meio modificado que continha CMC ou xilana como única fonte de carbono, denominado meio TESTE.

Primeiramente foi feito um pré-inóculo (com 5 mL do meio TESTE) com as estirpes SP245 e FP2, deixando-as incubadas por 3 dias. Após este período, passou-se todo o volume da cultura para um novo frasco contendo 100 mL de meio TESTE com antibióticos. Para avaliar o crescimento das bactérias, as estirpes foram inoculadas no meio TESTE em 6 diferentes formulações (maiores detalhes: item 4.2 em Material e Métodos):

- Meio TESTE com extrato de levedura (0,5 g/L), sem CMC ou xilana;
- Meio TESTE com extrato de levedura (0,5 g/L) com CMC (5 g/L)

- c) Meio TESTE com extrato de levedura (0,5 g/L) com xilana (3 g/L);
- d) Meio TESTE sem extrato de levedura e sem CMC ou xilana;
- e) Meio TESTE sem extrato de levedura com CMC (5 g/L);
- f) Meio TESTE sem extrato de levedura com xilana (3 g/L);

Os resultados obtidos estão mostrados na tabela 13.

Tabela 13: Cultivo de *A. brasilense* SP245 e FP2 na presença de diferentes polissacarídeos como fonte de carbono.

Cultivo em meio TESTE		Crescimento – SP245	Crescimento – FP2
Com extrato de levedura	-C	+	+
	+CMC	+	+
	+Xilana	+	+
Sem extrato de levedura	-C	-	-
	+CMC	-	-
	+Xilana	-	+

-C: sem adição de fonte de carbono. + : crescimento observado; – : crescimento não observado.

As duas estirpes de *A. brasilense* cresceram em todas as condições em que extrato de levedura estava presente, mesmo na presença de CMC ou xilana, sugerindo que o extrato de levedura (a 0,5 g/L) contém quantidade suficiente de fonte de carbono para sustentar o crescimento das bactérias.

A estirpe Sp245 não pode ser cultivada nos meios sem extrato de levedura. Já a estirpe FP2, apresentou crescimento no meio contendo xilana. Para avaliar se a cultura estava contaminada, uma alíquota do cultivo foi plaqueada em meio NFbHPN-lactato (sem antibióticos). Todas as colônias recuperadas eram de *A. brasilense* FP2.

O crescimento identificado demonstra que a estirpe FP2 tem maquinaria enzimática capaz de degradar a xilana, conseguindo utilizá-la em seu metabolismo. Isso pode indicar que as xilanases selecionadas neste trabalho possam ter atividade hidrolítica caso sejam expressas na forma nativa (sem cauda de histidina ou sem mutações) ou que sejam encontradas condições ótimas de atividade.

6 CONCLUSÃO

O genoma de *Azospirillum brasilense* apresenta 15 genes para glicosil-hidrolases, sendo que 10 deles foram escolhidos para serem estudados neste trabalho. Apenas 4 deles foram clonados, mas as proteínas expressas não apresentaram atividade em diferentes condições de reação. Em três das sequências clonadas foram identificadas mutação devidas a PCR, o que pode justificar a falta de atividade. É ainda possível que a falta de atividade pode ser devida a ausência de algum co-fator ou ainda que o substrato utilizado não seja o adequado.

Cultivo da estirpe FP2 de *A. brasilense* em meio contendo xilana revelou que esta estirpe é capaz de usar xilana como única fonte de carbono, indicando que esta estirpe pode apresentar xilanases ativas em seu metabolismo.

REFERÊNCIAS

- ALTSCHUL, S. F.; MADDEN, T. L.; SCHAFFER, A. A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic. Acids Res.**, v. 25, p. 3389-3402, 1997.
- AMORE, A.; PEPE, O.; VENTORINO, V.; BIROLO, L.; GIANGRANDE, C.; FARACO, V. Cloning and recombinant expression of a cellulase from the cellulolytic strain *Streptomyces* sp. G12 isolated from compost. **Microb. Cell Factories**, v. 11, p. 164-176, 2012.
- ARAI, T.; ARAKI, R.; TANAKA, A.; KARITA, S.; KIMURA, T.; SAKKA, K.; OHMIYA, K. Characterization of a cellulase containing a family 30 carbohydrate-binding module (CBM) derived from *Clostridium thermocellum* CelJ: importance of the CBM to cellulose hydrolysis. **Journal of Bacteriology**, v. 185, p. 504-512, 2003.
- BABCOCK, L. W. **U.S. Patent**, v.1, p.825-964, 1932.
- BALDANI, V.L.D.; BALDANI, J.I.; DÖBEREINER, J. Effect of *Azospirillum* inoculation on root infection and nitrogen incorporation in wheat. **Can. J. Microbiol.**, v. 29, p. 924-929, 1983.
- BNDES, CGEE, FAO, CEPAL. **Bioetanol de cana-de-açúcar: Energia para o desenvolvimento sustentável – Resumo executivo**. 2008.
- BNDES, CGEE. Livro: **Bioetanol de cana-de-açúcar: Energia para o desenvolvimento sustentável**. Rio de Janeiro: BNDES, 2008.
- BON, E.P.S.; FERRARA, M.A. Bioethanol production via enzymatic hydrolysis of cellulosic biomass. Disponível em: < <http://www.fao.org/biotech/docs/bon.pdf>>. Acessado em: 22/08/2012.
- BRADFORD, M. M. A rapid and sensitive method for the quantification of microgram quantities of protein utilization: the principle of protein-dye binding. **Anal. Biochem.**, v.72, p. 248-254, 1976.
- CHENG, J., SUN, Y. Hydrolysis of lignocellulosic materials for ethanol production: a review. **Bioresource Technology**, v. 83, p. 1-11, 2001.
- COUGHLAN, M. P.; LJUNGDAHL, L. G. Comparative biochemistry of fungal and bacterial cellulolytic enzyme system. In: AUBERT, J. P.; BEGUIN, P.; MILLET, J. **Biochemistry and Genetics of Cellulose Degradation**, p. 11-30, 1988.
- DE BUCK, E.; LAMMERTYN, E.; ANNÉ, J. The importance of the twin-arginine translocation pathway for bacterial virulence. **Trends Microbiol.**, v. 16, n.9, p. 442-453, 2008.
- FOODY, B. E.; FOODY, K. J. Development of an integrated system for producing ethanol from biomass. **Energy Biom. Wastes**, v. 14, p. 1225-1243, 1991.

FRACKMAN, S.; KOBBS, G.; SIMPSON, D.; STORTS, D. Betaine and DMSO: enhancing agents for PCR. **Promega Notes**, n. 65, p. 27-31, 1998.

GALANTE, Y.M.; DE CONTI, A.; MONTEVERDI, R. Application of *Trichoderma* enzymes in food and feed industries. In: HARMAN, G.F.; KUBICEK, C.P. **Trichoderma & Gliocladium – enzymes, biological control and commercial applications**. V. 2, Londres: Taylor & Francis, p. 327-342, 1998.

GASTEIGER, E.; HOOGLAND, C.; GATTIKER, A.; DUVAUD, S.; WILKINS, M.R.; APPEL, R.D.; BAIROCH, A. Protein identification and analysis tools on the ExPASy server. In WALKER, J.M. **The Proteomics Protocols Handbook**, Humana Press, p. 571-607, 2005.

HANKIN, L.; ANAGNOSTAKIS, L. Solid media containing carboxymethylcellulose to detect C_x cellulase activity of micro-organisms. **J. of Gen. Microbiol.**, v. 98, p. 109-115, 1976.

HAQQI, T.; ZHAO, X.; PANCIU, A.; YADAV, S.P. Sequencing in the presence of betaine: improvement in sequencing of the localized repeat sequence regions. **J. of Biomol. Tech.**, v. 13, p. 265-271, 2002.

HENRISSAT, B. A classification of glycosyl hydrolases based on amino acid sequence similarities. **Biochem. J.**, v. 280, p. 309-316, 1991.

HUERGO, L. F. Regulação do metabolismo de nitrogênio em *A. brasilense*. Curitiba, 2006. 170 f. Tese (Doutorado em Bioquímica) – Setor de Ciências Biológicas, Universidade Federal do Paraná.

JAIN, D. K. e PATRIQUIN, D. G. Root hair deformation, bacterial attachment, and plant growth in wheat-*azospirillum* associations. **App. and Env. Microb.**, v. 48, p. 1208-1213, 1984.

JONES, D.T. Protein secondary structure prediction based on position-specific scoring matrices. **J. Mol. Biol.**, v.292, p.195-202, 1999.

KOCHER, T. D.; WILSON, A. C. DNA amplification by the polymerase chain reaction. In BROWN, T. A., **Essential Molecular Biology: A Practical Approach**, IRL Press, Oxford University Press, Oxford, p.185-208, 1991.

KOK, L. IB Biology, enzyme kinetics, sucrose hydrolysis, glucose assay using colorimeter. Disponível em: <<http://lawrencekok.blogspot.com.br/2011/06/ib-biology-enzyme-kinetics-sucrose.html#sthash.aZ1wZ74q.dpuf>>. Acesso em: 29/11/2013.

LAEMMLI, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T7. **Nature**. v. 227, p. 680-685, 1970.

LARKIN, M.A.; BLACKSHIELDS, G.; BROWN, N.P.; CHENNA, R.; MCGETTIGAN, P.A.; MCWILLIAM, H.; VALENTIN, F.; WALLACE, I.M.; WILM, A.; LOPEZ, R.; THOMPSON, J.D.; GIBSON, T.J.; HIGGINS, D.G. ClustalW and ClustalX version 2. **Bioinfo.**, v. 23 (21), p. 2947-2948, 2007.

MACHADO, H.B.; YATES, M.G.; FUNAYAMA, S.; RIGO, L.U.; STEFFENS, M.B.R.; SOUZA, E.M.; PEDROSA, F.O. The *ntrBC* genes of *Azospirillum brasilense* are part of *nifR3*-like-*ntrB*-*ntrC* operon and are negatively regulated. **Can. J. Microbiol.**, v. 41, p.674-684, 1995.

MAMMEDOV, T.G.; PIENAAR, E.; WHITNEY, S.E.; TERMAAT, J.R.; CARVILL, G.; GOLIATH, R.; SUBRAMANIAN, A.; VILJOEN, H.J. A fundamental study of the PCR amplification of GC-rich DNA templates. **Comput. Biol. Chem.**, v. 32, n.6, p. 452-457, 2008.

MARCHLER-BAUER, A.; LU, S.; ANDERSON, J.B.; CHITSAZ, F.; DERBYSHIRE, M.K.; DEWEESE-SCOTT, C.; FONG, J.H.; GEER, L.Y.; GEER, R.C.; GONZALES, N.R.; GWADZ, M.; HURWITZ, D.I.; JACKSON, J.D.; KE, Z.; LANCZYCKI, C.J.; LU, F.; MARCHLER, G.H.; MULLOKANDOV, M.; OMELCHENKO, M.V.; ROBERTSON, C.L.; SONG, J.S.; THANKI, N.; YAMASHITA, R.A.; ZHANG, D.; ZHENG, C.; BRYANT, S.H. CDD: a conserved domain database for the functional annotation of proteins. **Nucleic Acids Res.**, v. 39, p. 225-229, 2010.

MEHDIPOUR-MOGHADDAM, M. J.; GITI, E.; BOUZARI, M.; MOSTAJERAN, A.; SALEHI, Z. Novel phytase and cellulase activities in endophytic azospirilla. **World App. Sci. J.**, v. 10, p. 1129-1135, 2010.

MIETTINEN-OINONEN, A. PALOHEIMO, M.; LANTTO, R.; SUOMINEN, P. Enhanced production of cellobiohydrolases in *Trichoderma reesei* and evaluation of the new preparations in biofinishing of cotton. **J. of Biotech.**, v. 116 (3), p. 305-317, 2005.

MIETTINEN-OINONEN, A. *Trichoderma reesei* strains for production of cellulases for the textile industry. **VTT publications**, v. 550, 2004.

MILLER, G.L., Use of dinitrosalicylic acid reagent for determination of reducing sugar. **Anal. Chem.**, n. 31, p.426, 1959.

NAKAI, K.; KANEHISA, M. Expest system for predicting protein localization sites in gram-negative bacteria. **Proteins**, v. 11, p. 95-110, 1991.

NOVOZYMES – artigo: **Pode o branco ser mais branco?** Araucária: HC, 2003.

PANG, H., ZHANG, P., DUAN, C., MO, X., TANG, J., FENG, J. Identification of cellulase genes from the metagenomes of compost soils and functional characterization of one novel endoglucanase. **Curr. Microbiol.**, v. 58, p. 404–408, 2009.

PEDROSA, F.O.; YATES, M.G. Regulation of nitrogen fixation (*nif*) genes of *Azospirillum brasilense* by *nifA* and *ntrC* (*glnG*) type genes. **FEMS Microbiol. Lett.**, v. 55, p. 95-101, 1984.

SAMBROOK, J.; FRITSCH, E.F.; MANIATIS, T. **Molecular cloning a laboratory manual**. 2 ed. New York. Cold Spring Harbor Laboratory Press, 1989.

SANGER, F.; NICKLEN, S.; COULSON, A. R. DNA sequencing with chain-termination inhibitors. **Proc. Natl. Acad. Sci. USA.**, v. 74, p. 5463-5467, 1977.

SCHELLER, H. V.; ULVSKOV, P. Hemicelluloses. **Annu. Rev. Plant Biol.**, v. 61, p. 263-289, 2010.

SHEVCHENKO, A.; WILM, M.; VORM, O.; e MANN, M. Mass spectrometric sequencing of proteins from silver-stained polyacrilamide gels. **Analytical Chemistry**, v. 68, p. 850-858, 1996.

SIGRIST C.J.A.; CERUTTI L.; HULO N.; GATTIKER A.; FALQUET L.; PAGNI M, BAIROCH A, BUCHER P. PROSITE: a documented database using patterns and profiles as motif descriptors. **Brief Bioinform.**, v. 3, p. 265-274, 2002.

SOUZA, E.M. Clonagem, caracterização e seqüenciamento dos genes nifA e nifB de *Herbaspirillum seropedicae*. **Tese de Doutorado**. Programa de Pós-Graduação em Ciências-Bioquímica. Departamento de Bioquímica e Biologia Molecular – UFPR, Curitiba, p. 260, 1990.

STEENHOUDT, O.; VANDERLEYDEN, J. *Azospirillum*, a free-living nitrogen-fixing bacterium closely associated with grasses: genetic, biochemical and ecological aspects. **FEMS Microbiol. Rev.**, v. 24, p. 487–506, 2000.

TAHERZADEH, M.J.; KARIMI, K. Enzyme-based hydrolysis processes for ethanol from lignocellulosic materials: a review. **Biores.**, v. 2(4), p. 707-738, 2007.

WISNIEWSKI-DYE, F., BORZIAK, K., KHALSA-MOYERS, G., ALEXANDRE, G., SUKHARNIKOV, L. O., WUICHET, K., HURST, G. B., MCDONALD, W. H., ROBERTSON, J. S., BARBE, V., CALTEAU, A., ROUY, Z., MANGENOT, S., PRIGENT-COMBARET, C., NORMAND, P., BOYER, M. L., SIGUIER, P., DESSAUX, Y., ELMERICH, C., CONDEMINE, G., KRISHNEN, G., KENNEDY, I., PATERSON, A. H., GONZALEZ, V., MAVINGUI, P., ZHULIN, I. B. *Azospirillum* genomes reveal transition of bacteria from aquatic to terrestrial environments. **Plos genetics**, v. 7, 2011.

WOOD, T.M.; BHAT, K. M. Methods for measuring cellulase activities. **Methods Enzymol.**, v. 160, p. 87–117, 1988.

YAN, X.; GENG, A.; ZHANG, J.; WEI, Y.; ZHANG, L.; QIAN, C.; WANG, Q.; WANG, S.; ZHOU, Z. Discovery of (hemi-) cellulase genes in a metagenomic library from a biogas digester using 454 pyrosequencing. **Appl. Microbiol. Biotechnol.**, v. 97, p. 8173-8182, 2013.

YOUNG, J. P. W. Phylogenetic classification of nitrogen-fixing organisms. In: STACEY, G.; BURRIS, R. H.; EVANS, H. J. (eds.). **Biological Nitrogen Fixation**. London: Chapman Hall, New York, p.43-86, 1992.

ZHANG, Y., LYND, L.R. A functionally based model for hydrolysis of cellulose by fungal cellulase. **Biotech. Bioeng.**, v. 94, n. 5, p. 888–898, 2006.